REPRODUCTIVE PLASTICITY AND THE EVOLUTION OF THE INSECT SOCIETIES

BY

BERYL M. JONES

DISSERTATION

Submitted in partial fulfillment of the requirements
for the degree of Doctor of Philosophy in Ecology, Evolution, and Conservation Biology
in the Graduate College of the
University of Illinois at Urbana-Champaign, 2019

Urbana, Illinois

Doctoral Committee:

Professor Gene E. Robinson, Chair and Director of Research
Professor Alison M. Bell
Professor Matthew E. Hudson
Professor Andrew V. Suarez

ABSTRACT

A fundamental goal of evolutionary biology is to understand how novel traits arise. Eusociality represents an extreme form of social organization which has evolved independently a number of times across insects and is characterized especially in the Hymenoptera by a novel polyphenism between reproductive (queen) and non-reproductive (worker) castes. While a growing body of research continues to improve our understanding of the mechanisms underlying the development of these castes, less is known about how castes evolved from solitary ancestors. In this dissertation, I leverage naturally-occurring social plasticity in two species of bees to shed light on potential mechanisms of caste evolution across social insects. In Chapter 1, I provide a detailed overview of the work contained within this dissertation. In Chapter 2, I develop a perspective on how ancestral behavioral plasticity may have facilitated the evolution of castes through genetic accommodation. In Chapter 3, I present a *de novo* transcriptome assembly for *Megalopta genalis,* a facultatively eusocial sweat bee that exhibits multiple social phenotypes within one population and may therefore represent a transition between solitary and social reproduction. I use this transcriptome in Chapter 4 to identify gene expression differences associated with social phenotypes of *M. genalis,* and compare these to genes involved in caste determination of other eusocial species as well as genes implicated in the evolution of eusociality through comparative studies of bees. In Chapter 5, I use a high-resolution behavioral tracking system to discover a previously undescribed form of colony organization in honey bees that occurs after a colony loses and is unable to replace its queen and some workers begin to lay eggs. Surprisingly similar to the social variation observed across nests of *M. genalis,* these colonies of honey bee workers display multiple levels of social plasticity, evoking transitional stages in eusocial evolution associated with the venerable Ovarian Ground Plan Hypothesis. Finally, in Chapter 6, I use transcriptomics and chromatin accessibility analyses of bees in laying worker colonies to explore how changes in brain gene regulation may contribute to variation in colony social organization, with comparative analyses to place this variation in the broader context of caste evolution across social insect lineages.

ACKNOWLEDGMENTS

I was fortunate to be surrounded by many wonderful scientists and friends throughout my graduate career, and I owe thanks to many people for making this dissertation possible. My advisor, Gene Robinson, encourages a spirit of discovery and scholarship that permeates his research group, and his leadership has fostered academic friendships that will last a lifetime. Gene empowered me to follow the trails of exciting results and to forge my own path in research, a freedom for which I will always be grateful. I could not have asked for a better mentor for my Ph.D., and I hope to live up to the example he has set for me.

I would also like to thank Andy Suarez, Alison Bell, and Matt Hudson for serving on my committee. They each provided valuable feedback and guidance as I navigated through my time as a graduate student, with unique perspectives that shaped and improved my research. Countless other professors at UIUC are to thank for creating a wonderful atmosphere of collegiality and an exceptional environment in which to mature as a scientist.

Members of the Robinson lab were wonderful colleagues who provided ideas, hands, and helpful conversations throughout my Ph.D. They were all important in shaping my research and experience in the lab, but a few deserve special mention here. Amy Cash Ahmed was instrumental in guiding me through the molecular aspects of my work, giving me a set of skills I will carry with me always. Amy also provided emotional support when I faced the biggest challenges in my work, and continues to be a close friend. Charley Nye and Alison Sankey provided expert beekeeping support, and are more valuable than they know. Tim Gernat created an amazing set of research tools and helped me to use them in exciting new ways that I never thought possible. Claudia Lutz looked over drafts of manuscripts, talked research ideas, and helped translate my research to the public. Allyson Ray, Lindsey Block, Sarah Bransley, Jeannette Cullum, and so many other students and technicians worked tirelessly on demanding research projects, and without them much of this dissertation would not have happened. Last but certainly not least, Ian Traniello, with whom I share so many unique bonds and who will always understand me in ways others cannot- I am proud to be your friend.

A number of other collaborators and friends have guided and encouraged me. Bill Wcislo welcomed me to Panama and showed me the wonder of the tropics. Callum Kingwell shared many early mornings with me filming *Megalopta* nests, and was always generous with his time.

Vikyath Rao tolerated and obliged many late night requests for data and figures, always with a smile. Karen Kapheim continually provided examples of scientific excellence; I will always strive to follow her lead and am proud to be her collaborator. And to my friend Alex Walton, who keeps me company in my moves further and further east- may we both end up back where we started, working together under the Tucson sun.

Finally, I give thanks to members of my family, who have always been supportive of my goals, even when they may not have understood what those goals were. I owe much of my success to the support and encouragement of my best friend and partner, Troy. Thank you for your constant support and strength; I will forever be grateful for you.

TABLE OF CONTENTS

CHAPTER 1

INTRODUCTION

The evolution of eusociality represents a major transition in life (Smith and Szathmaáry, 1995) that has occurred at least 24 times independently (Bourke, 2011) and has produced some of the most ecologically successful taxa on the planet (Wilson, 1985). Biologists have long been interested in describing the evolutionary mechanisms involved in solitary to eusocial transitions, both for their specific scientific merits and to understand phenotypic novelty more broadly.

A defining feature of eusociality is the presence of queens and workers, who perform markedly different suites of behavioral tasks in the colony despite arising from the same genomic background, a polyphenism driven by nutritional and environmental inputs. This polyphenism has been well studied in the context of differentially expressed genes during development, as well as downstream effects of these developmental differences (e.g., Feldmeyer et al., 2014; Grozinger et al., 2007; Harrison et al., 2015; He et al., 2017; Jones et al., 2017). However, the processes of evolution leading to selection for these bifurcating developmental pathways are not well understood.

The hymenopteran insect order, which contains the ants, bees and wasps, features the greatest number of independent evolutions of eusociality, currently estimated to be 9 (Bourke, 2011; Cruz, 1981; Gibbs et al., 2012; Hines et al., 2007; Hughes et al., 2008; Moreau et al., 2006; Romiguier et al., 2016; Schwarz et al., 2007). Bees are ideal organisms in which to explore the origins of caste-related phenotypic plasticity. Extant bees span a range of social phenotypes and are nested within a broader phylogeny including both solitary and social members. This naturally occurring variation in social complexity has allowed for comparative genomic analyses to identify patterns of evolution associated with social behavior across bee lineages (Kapheim et al., 2015; Woodard et al., 2011). In addition, individuals in some species of bees display flexible social strategies, allowing for studies of the mechanisms underlying social variation (Jones et al., 2017; Kocher and Paxton, 2014; Kocher et al., 2018). Genomic resources for bees are also increasing steadily, with more than 15 publicly available genomes across multiple lineages and independent origins of eusociality in bees (Kapheim et al., 2015; Kapheim et al., 2019; Kocher et al., 2013; Sadd et al., 2015). Armed with these resources, it is now more possible than ever to

explore the mechanistic basis of caste evolution across independent origins of eusociality.

A number of leading hypotheses regarding the origins of queen and worker castes involve changes in the timing and regulation of ancestral gene regulatory networks (Linksvayer and Wade, 2005; West-Eberhard, 1987; West-Eberhard, 1996). The first of these, the Ovarian Ground Plan Hypothesis (OGPH), suggests that an ancestral solitary cycle of reproductive and non-reproductive behaviors was decoupled to delineate the queen and worker castes of eusocial species through gene regulatory changes acting on ancestral plasticity (West-Eberhard, 1987; West-Eberhard, 1996). Hypotheses like the OGPH stress the importance of ancestral phenotypic plasticity in the origin of castes, rather than novel mutations or novel genes in transitions to eusocial behavior (Ferreira et al., 2013; Sumner, 2013). The role of phenotypic plasticity in evolution has remained controversial, but "plasticity-first" models of evolution have been proposed for over a century, when Baldwin and others suggested that fitness differences arising from phenotypic plasticity during development could lead to genetic change over many generations through a process now known as genetic accommodation (Baldwin, 1896; Baldwin, 1902; Morgan, 1896).

In Chapter 2, I review how ancestral plasticity in behavior may have facilitated the evolution of caste in eusocial bees and other taxa through genetic accommodation. While mechanisms of genetic accommodation are not well understood, advances in the field of transgenerational epigenetic inheritance are bringing us closer than ever to bridging gaps between theory and empirical data. I review the literature on transgenerational epigenetic inheritance, as well as traits of social insects that make them particularly well suited for studies of genetic accommodation. I also discuss current literature consistent with a role for genetic accommodation in the evolution of social insect castes. Finally, I propose a novel empirical method called 'eusocial engineering,' which involves environmental induction of novel social phenotypes followed by close mechanistic studies to understand how environmentally induced plasticity may lead to heritable changes in social behavior. This chapter has been published in the *Journal of Experimental Biology* (Jones and Robinson, 2018).

Chapter 3 discusses the development of a *de novo* transcriptome assembly for a bee with polymorphic behavioral phenotypes. *Megalopta genalis* is a tropical sweat bee in the family Halictidae that is facultatively eusocial and displays a wide range of social behaviors within a single population (Wcislo and Gonzalez, 2006). The halictid lineage contains multiple origins of

eusociality (Brady et al., 2006; Gibbs et al., 2012), and phylogenetic studies of bees point to a solitary ancestral lifestyle (Wilson 1971), suggesting that *M. genalis* may represent a transition between solitary and social reproduction. In addition to displaying both solitary and social phenotypes within the same population, *M. genalis* adults are also remarkably flexible in their reproductive ability. In social nests, queen loss results in a transition to egg-laying in previously non-reproductive workers (Smith et al. 2009). These replacement queens subsequently achieve the same egg-laying rate as a solitary reproductive female (Smith et al. 2009), suggesting no physiological limitation to reproduction in *M. genalis* workers. This behavioral flexibility is absent in some more advanced forms of eusociality, but may have been an existing trait in the ancestors of eusocial lineages. I use the assembled transcriptome to assess developmental and sex-related differences in gene expression in this socially polymorphic bee, providing the first available transcriptomic resource for a bee in the family Halictidae. This chapter has been published in *G3:Genes|Genomes|Genetics* (Jones et al., 2015).

Chapter 4 leverages the transcriptome assembly created in Chapter 3 to identify differentially expressed genes associated with different social phenotypes of *M. genalis.* I first performed field work to behaviorally characterize nests so I could evaluate gene expression differences in two tissues (brain and abdomen) between solitary reproductives, queens, and workers. I also performed a queen-removal manipulation in social nests and assessed whether replacement queens (workers who shifted to reproductive behavior following queen-removal) exhibited gene expression profiles shifted toward that of reproductive phenotypes. Caste differences were much stronger in tissues of the abdomen compared with brain; abdominal tissues showed a clear signature of reproductive activity and a strong shift in the gene expression profiles of replacement queens.

In addition to finding a strong transcriptomic signature of reproductive plasticity in *M. genalis,* I also discovered significant overlap between worker-biased genes (genes more highly expressed in workers relative to queens) in *M. genalis* and those undergoing selection in obligately eusocial lineages of bees. This finding was particularly pronounced for genes in the glycolysis pathway, which has been previously implicated in the evolution of caste in multiple social insect species (Woodard et al. 2011, Berens et al. 2015). These results suggest a role for genetic accommodation in caste evolution, with ancestral transcriptomic plasticity providing a substrate for selection in the evolution of worker castes. Also striking was the strong overlap

between *M. genalis* worker-biased genes and those undergoing selection in the honey bee, *Apis mellifera* (Harpur et al. 2014), suggesting common pathways to worker behavior despite ~115 million years of divergence (Cardinal and Danforth 2013). Chapter 4 has been published in *Proceedings of the Royal Society B* (Jones et al., 2017).

My work in *M. genalis* suggests that, consistent with the ideas presented in Chapter 2, ancestral transcriptomic plasticity may have been involved in the evolution of caste. However, the lack of genomic resources and inability to deeply probe transcriptomic and epigenomic signatures of behavior in this species are obstacles to achieving a strong mechanistic understanding of eusocial evolution. The honey bee, *Apis mellifera*, lacks the facultative eusociality that makes *M. genalis* a strong model system for understanding evolutionary origins of caste, but has the advantages of a well-annotated genome and development of many tools related to gene and behavioral manipulation. In addition, a recent finding about honey bee colonies with egg-laying workers (LW colonies), which occur when a colony becomes permanently queenless, suggested that LWs have more extensive behavioral plasticity than previously appreciated. Naeger et al. (2013) discovered that in LW colonies, many workers with activated ovaries perform typical non-reproductive behaviors such as foraging and defending the hive. This potential coupling of reproductive and non-reproductive tasks in single honey bee individuals is reminiscent of ancestral forms of social behavior, and may allow for a mechanistic understanding of how caste-related behaviors were decoupled in evolution. Before LWs can be used as models to understand mechanisms of this decoupling, however, it was necessary to confirm that LW engaged in both egg-laying and non-reproductive behaviors, since Naeger et al. (2013) primarily relied on ovarian development as a proxy for reproduction. In addition, the extent of interindividual plasticity among LW was unknown, as was the overall social organization of LW colonies.

In Chapter 5, I employed a high-resolution behavioral tracking system (Gernat et al., 2018) to obtain behavioral data on every bee in six colonies of LW honey bees. This barcode-based tracking system uses one-second-resolution images and convolutional neural networks to automatically detect bees performing three behaviors: egg-laying (a reproductive behavior), foraging (an important and commonly studied non-reproductive behavior), and trophallaxis (a social behavior between a pair of bees, where food and putative communication-related molecules are transferred; Free, 1956; Leboeuf et al., 2016; Nixon and Ribbands, 1952). Using

these behavioral data, I discovered a striking pattern of social organization in colonies of LWs relative to queenright colonies. In LW colonies, a subset of individuals perform both egg-laying and foraging behaviors, much like the presumed solitary ancestor to honey bees (Michener, 1974; Wilson, 1971). These "generalist" bees confirmed the ability of LW to engage in non-reproductive behaviors with developed ovaries, as previously discovered (Naeger et al., 2013), while cohabiting a nest with other individuals similar to communal species of bees and wasps (e.g., Abrams and Eickwort, 1981; Danforth, 1989; McCorquodale and Naumann, 1988). However, generalist bees were less common than bees that specialized on either egg-laying or foraging tasks, participating in a division of labor similar to that seen in primitively eusocial species of bees and wasps (Donnell, 1998; Michener, 1974; Turillazzi and West-Eberhard, 1996). Bees engaged in this reproductive division of labor fell along a gradient of behavioral specialization, with high skew in task performance and the presence of both highly specialized layers and foragers. Generalist bees, as well as specialized layers and foragers, were engaged in a social network with properties similar to those of queenright colonies, suggesting that LW colonies are a coordinated alternative social state of honey bees. The unique social organization of LW colonies, with the co-occurrence of communal and primitively eusocial individuals, suggests that LW colonies may provide a glimpse into transitional states in social behavior. Communal generalists may have been precursors to primitive eusociality, representing stable intermediate stages during the evolution of eusociality from a solitary ancestor (Evans and West-Eberhard, 1970; Rehan and Toth, 2015). LW colonies thus represent a unique opportunity to study the molecular underpinnings of socio-behavioral plasticity, with implications for understanding potential evolutionary transitions in sociality.

In Chapter 6, I leverage the behavioral variation present in LW colonies and sequenced brain mRNA and accessible brain chromatin of highly specialized egg layers, specialized foragers, and generalists to explore how flexibility in gene regulation may influence this evolutionarily relevant variation in social behavior. I found that these three groups of bees displayed remarkably distinct brain transcriptomic profiles, as well as differences in chromatin accessibility. Genes which differentiated specialized layers and foragers were enriched for genes associated with reproductive castes in other social insect species. In addition, brain gene expression differences between these three groups resembled those previously found to be associated with foragers and nurses from queenright colonies, consistent with hypotheses

suggesting that ancestral reproductive signaling pathways were co-opted during the evolution of worker-related behaviors, an extension of the OGPH (Amdam et al., 2008; Graham et al., 2011; Page et al., 2012). Foragers in LW colonies also showed strong upregulation of metabolic pathways in the brain, including glycolysis, which I also found to be enriched among worker-related genes in *Megalopta genalis* in Chapter 4. The glycolysis pathway has also been implicated in the evolution of caste in multiple social insect species (Woodard et al. 2011, Berens et al. 2015). Together, these results suggest that LW behavioral plasticity may be regulated by ancestral gene regulatory networks that were also important in the evolution of reproductive castes in bees and other social insects.

While the brain transcriptomic profiles of egg-layers and foragers within LW colonies were highly distinct, I found that generalist bees that performed both egg-laying and foraging were intermediate both in their transcriptomic as well as chromatin accessibility landscapes. In addition, brain gene expression and chromatin accessibility showed continuous variation across the three groups, and moreover, it was correlated with the continuous behavioral variation observed across individuals in LW colonies. These results suggest that the seemingly discretized behavioral states of the honey bee worker, including layers and foragers within a LW colony, are mediated through gene regulatory networks that are highly flexible and continuous. Together with the continuous behavioral variation observed in Chapter 5, these results suggest that variation in social organization may emerge through incremental changes in gene regulatory networks from an ancestrally plastic and continuous distribution of phenotypes.

Studies of social insects have not only led to exciting discoveries related to the mechanisms underlying social behavior, but have also contributed to our growing understanding of how phenotypic plasticity is mediated. In this dissertation, I present my own contributions to this topic, with the goal of using techniques from genomics to advance our understanding of eusocial evolution in the broader context of how phenotypic plasticity may shape and be shaped by evolutionary processes.

References

Abrams, J. and Eickwort, G. C. (1981). Nest switching and guarding by the communal sweat bee *Agapostemon virescens* (Hymenoptera, Halictidae). *Insectes Soc.* 28, 105–116.

Amdam, G. V, Page Jr, R. E. and Tsuruda, J. M. (2008). Sensory response system of social behavior tied to female reproductive traits. *PLoS One* 3, e3397.

Baldwin, J. M. (1896). A new factor in evolution. *Am. Nat.* 30, 441–451.

Baldwin, J. M. (1902). *Development and Evolution*. London: MacMillan Co.

Bourke, A. F. G. (2011). *Principles of Social Evolution*. Oxford: Oxford University Press.

Brady, S. G., Sipes, S., Pearson, A. and Danforth, B. N. (2006). Recent and simultaneous origins of eusociality in halictid bees. *Proc. R. Soc. B Biol. Sci.* 273, 1643–1649.

Cruz, Y. P. (1981). A sterile defender morph in a polyembryonic hymenopterous parasite. *Nature* 294, 446–447.

Danforth, B. N. (1989). Nesting behavior of four species of *Perdita* (Hymenoptera: Andrenidae). *J. Kansas Entomol. Soc.* 62, 59–79.

Donnell, S. O. (1998). Reproductive caste determination in eusocial wasps (Hymenoptera: Vespidae). *Annu. Rev. Entomol.* 43, 323–346.

Evans, H. and West-Eberhard, M. (1970). *The Wasps*. Ann Arbor: University of Michigan Press.

Feldmeyer, B., Elsner, D. and Foitzik, S. (2014). Gene expression patterns associated with caste and reproductive status in ants: Worker-specific genes are more derived than queen-specific ones. *Mol. Ecol.* 23, 151–161.

Ferreira, P. G., Patalano, S., Chauhan, R., Ffrench-Constant, R., Gabaldón, T., Guigó, R. and Sumner, S. (2013). Transcriptome analyses of primitively eusocial wasps reveal novel insights into the evolution of sociality and the origin of alternative phenotypes. *Genome Biol.* 14, R20.

Free, J. (1956). A study of the stimuli which release the food begging and offering responses of worker honeybees. *Br. J. Anim. Behav.* 4, 95–101.

Gernat, T., Rao, V. D., Middendorf, M., Dankowicz, H., Goldenfeld, N. and Robinson, G. E. (2018). Automated monitoring of behavior reveals bursty interaction patterns and rapid spreading dynamics in honeybee social networks. *Proc. Natl. Acad. Sci.* 115, 1433–1438.

Gibbs, J., Brady, S. G., Kanda, K. and Danforth, B. N. (2012). Phylogeny of halictine bees supports a shared origin of eusociality for *Halictus* and *Lasioglossum* (Apoidea: Anthophila: Halictidae). *Mol. Phylogenet. Evol.* 65, 926–939.

Graham, A. M., Munday, M. D., Kaftanoglu, O., Page, R. E., Amdam, G. V. and Rueppell, O. (2011). Support for the reproductive ground plan hypothesis of social evolution and major QTL for ovary traits of Africanized worker honey bees (*Apis mellifera* L.). *BMC Evol. Biol.* 11, 95.

Grozinger, C. M., Fan, Y., Hoover, S. E. R. and Winston, M. L. (2007). Genome-wide analysis reveals differences in brain gene expression patterns associated with caste and reproductive status in honey bees (*Apis mellifera*). *Mol. Ecol.* 16, 4837–4848.

Harrison, M. C., Hammond, R. L. and Mallon, E. B. (2015). Reproductive workers show queenlike gene expression in an intermediately eusocial insect, the buff-tailed bumble bee *Bombus terrestris*. *Mol. Ecol.* 24, 3043–3063.

He, X. J., Jiang, W. J., Zhou, M., Barron, A. B. and Zeng, Z. J. (2017). A comparison of honeybee (*Apis mellifera*) queen, worker and drone larvae by RNA-Seq. *Insect Sci.* 00, 1–11.

Hines, H. M., Hunt, J. H., O'Connor, T. K., Gillespie, J. J. and Cameron, S. A. (2007). Multigene phylogeny reveals eusociality evolved twice in vespid wasps. *Proc. Natl. Acad. Sci.* 104, 3295–3299.

Hughes, W. O. H., Oldroyd, B. P., Beekman, M. and Ratnieks, F. L. W. (2008). Ancestral monogamy shows kin selection is key to the evolution of eusociality. *Science* 320, 1213–1216.

Jones, B. M. and Robinson, G. E. (2018). Genetic accommodation and the role of ancestral plasticity in the evolution of insect eusociality. *J. Exp. Biol.* 221, jeb153163.

Jones, B. M., Wcislo, W. T. and Robinson, G. E. (2015). Developmental transcriptome for a facultatively eusocial bee, *Megalopta genalis*. *G3* 5, 2127.

Jones, B. M., Kingwell, C. J., Wcislo, W. T. and Robinson, G. E. (2017). Caste-biased gene expression in a facultatively eusocial bee suggests a role for genetic accommodation in the evolution of eusociality. *Proc. R. Soc. B Biol. Sci.* 284, 20162228.

Kapheim, K. M., Pan, H., Li, C., Salzberg, S. L., Puiu, D., Magoc, T., Robertson, H.M., Hudson, M.E., Venkat, A., Fischman, B.J. et al. (2015). Genomic signatures of evolutionary transitions from solitary to group living. *Science* 348, 1139–1143.

Kapheim, K. M., Pan, H., Li, C., Blatti, C., Harpur, B. A., Ioannidis, P., Jones, B. M., Kent, C. F., Ruzzante, L., Sloofman, L., et al. (2019). Draft genome assembly and population genetics of an agricultural pollinator, the solitary alkali bee (Halictidae: *Nomia melanderi*). *G3* 9, 200865.2018.

Kocher, S. D. and Paxton, R. J. (2014). Comparative methods offer powerful insights into social evolution in bees. *Apidologie* 45, 289–305.

Kocher, S. D., Li, C., Yang, W., Tan, H., Yi, S. V., Yang, X., Hoekstra, H. E., Zhang, G., Pierce, N. E. and Yu, D. W. (2013). The draft genome of a socially polymorphic halictid bee, *Lasioglossum albipes*. *Genome Biol.* 16, 34.

Kocher, S. D., Mallarino, R., Rubin, B. E. R., Yu, D. W., Hoekstra, H. E. and Pierce, N. E. (2018). The genetic basis of a social polymorphism in halictid bees. *Nat. Commun.* 9, 4338.

Leboeuf, A. C., Waridel, P., Brent, C. S., Gonçalves, A. N., Menin, L., Ortiz, D., Riba-Grognuz, O., Koto, A., Soares, Z. G., Privman, E., et al. (2016). Oral transfer of chemical cues, growth proteins and hormones in social insects. *eLife* 5, e20375.

Linksvayer, T. A. and Wade, M. J. (2005). The evolutionary origin and elaboration of sociality in the aculeate Hymenoptera: maternal effects, sib-social effects, and heterochrony. *Q. Rev. Biol.* 80, 317–336.

McCorquodale, D. B. and Naumann, I. D. (1988). A new Australian species of communal ground nesting wasp, in the genus *Spilomena* Shuckard (Hymenoptera: Sphecidae: Pemphredoninae). *Aust. J. Entomol.* 27, 221–232.

Michener, C. D. (1974). *The Social Behavior of the Bees: A Comparative Study*. Cambridge, MA: Harvard University Press.

Moreau, C. S., Bell, C. D., Vila, R., Archibald, S. B. and Pierce, N. E. (2006). Phylogeny of the ants: Diversification in the age of angiosperms. *Science* 312, 101–104.

Morgan, C. (1896). Of modification and variation. *Science* 4, 733–740.

Naeger, N. L., Peso, M., Even, N., Barron, A. B. and Robinson, G. E. (2013). Altruistic behavior by egg-laying worker honeybees. *Curr. Biol.* 23, 1574–1578.

Nixon, H. L. and Ribbands, C. R. (1952). Food transmission within the honeybee community. *Proc. R. Soc. B Biol. Sci.* 140, 43–50.

Page, R. E., Rueppell, O. and Amdam, G. V. (2012). Genetics of reproduction and regulation of honeybee (*Apis mellifera* L.) social behavior. *Annu. Rev. Genet.* 46, 97–119.

Rehan, S. M. and Toth, A. L. (2015). Climbing the social ladder: the molecular evolution of sociality. *Trends Ecol. Evol.* 30, 426–433.

Romiguier, J., Cameron, S. A., Woodard, S. H., Fischman, B. J., Keller, L. and Praz, C. J. (2016). Phylogenomics controlling for base compositional bias reveals a single origin of eusociality in corbiculate bees. *Mol. Biol. Evol.* 33, 670–678.

Sadd, B. M., Barribeau, S. M., Bloch, G., de Graaf, D. C., Dearden, P., Elsik, C. G., Gadau, J., Grimmelikhuijzen, C. J. P., Hasselmann, M., Lozier, J. D., et al. (2015). The genomes of two key bumblebee species with primitive eusocial organization. *Genome Biol.* 16, 76.

Schwarz, M. P., Richards, M. H. and Danforth, B. N. (2007). Changing paradigms in insect social evolution: Insights from Halictine and Allodapine bees. *Annu. Rev. Entomol.* 52, 127–150.

Smith, J. M. and Szathmaáry, E. (1995). *The Major Transitions in Evolution*. Oxford: Oxford University Press.

Sumner, S. (2013). The importance of genomic novelty in social evolution. *Mol. Ecol.* 23, 26–28.

Turillazzi, S. and West-Eberhard, M. (1996). *Natural History and Evolution of Paper-Wasps*. Oxford: Oxford University Press.

Wcislo, W. T. and Gonzalez, V. H. (2006). Social and ecological contexts of trophallaxis in facultatively social sweat bees, *Megalopta genalis* and *M. ecuadoria* (Hymenoptera, Halictidae). *Insectes Soc.* 53, 220–225.

West-Eberhard, M. J. (1987). *Flexible strategy and social evolution*. (ed. Ito, Y., Brown, J. L., and Kikkawa, J.). Tokyo: Japan Scientific Societies Press.

West-Eberhard, M. J. (1996). Wasp societies as microcosms for the study of development and evolution. In *Natural history and evolution of paper wasps (eds Turillazzi S, West-Eberhard MJ)*, pp. 290–317.

Wilson, E. O. (1971). *The insect societies*. Cambridge: Harvard University Press.

Wilson, E. O. (1985). The sociogenesis of insect colonies. *Science* 228, 1489–95.

Woodard, S. H., Fischman, B. J., Venkat, A., Hudson, M. E., Varala, K., Cameron, S. A., Clark, A. G. and Robinson, G. E. (2011). Genes involved in convergent evolution of eusociality in bees. *Proc. Natl. Acad. Sci.* 108, 7472–7477.

CHAPTER 2

GENETIC ACCOMMODATION AND THE ROLE OF ANCESTRAL PLASTICITY IN THE EVOLUTION OF INSECT EUSOCIALITY[1]

Abstract

For over a century, biologists have proposed a role for phenotypic plasticity in evolution, providing an avenue for adaptation in addition to 'mutation-first' models of evolutionary change. According to the various versions of this idea, the ability of organisms to respond adaptively to their environment through phenotypic plasticity may lead to novel phenotypes that can be screened by natural selection. If these initially environmentally induced phenotypes increase fitness, then genetic accommodation can lead to allele frequency change, influencing the expression of those phenotypes. Despite the long history of 'plasticity-first' models, the importance of genetic accommodation in shaping evolutionary change has remained controversial – it is neither fully embraced nor completely discarded by most evolutionary biologists. We suggest that the lack of acceptance of genetic accommodation in some cases is related to a lack of information on its molecular mechanisms. However, recent reports of epigenetic transgenerational inheritance now provide a plausible mechanism through which genetic accommodation may act, and we review this research here. We also discuss current evidence supporting a role for genetic accommodation in the evolution of eusociality in social insects, which have long been models for studying the influence of the environment on phenotypic variation, and may be particularly good models for testing hypotheses related to genetic accommodation. Finally, we introduce 'eusocial engineering', a method by which novel social phenotypes are first induced by environmental modification and then studied mechanistically to understand how environmentally induced plasticity may lead to heritable changes in social behavior. We believe the time is right to incorporate genetic accommodation into models of the evolution of complex traits, armed with new molecular tools and a better understanding of non-genetic heritable elements.

---

[1] The work in this chapter has been adapted with permission from a previously published article: Jones, BM and Robinson, GE. 2018. Genetic accommodation and the role of ancestral plasticity in the evolution of insect eusociality. *Journal of Experimental Biology* 221:jeb153163.

Introduction

A compelling question in evolutionary biology involves the origins and evolution of novel traits. For centuries, biologists have been interested in the diversity of phenotypes across life and how this diversity arose. The role of genetic factors in the origination of novel traits has been especially well studied, including the roles of mutation, genetic drift and recombination in producing novel genetic combinations and phenotypes (Carroll, 2008). 'Mutation-first evolution' (see Glossary 2.1), where a new mutation provides novel phenotypes that can be 'screened' by natural selection, is easily studied when the mutation can be directly linked to the phenotype. Even without knowledge of the phenotypic consequences of alleles, mutation-first evolution studies can be initiated in both natural populations and laboratories simply by documenting changes in allele frequencies over time.

However, novel traits are also suggested to originate independent of new mutations, via the environmental and developmental induction of phenotypes. One of the first biologists to emphasize this was Baldwin, who at the turn of the 20th century suggested a process of 'organic selection' by which fitness differences arising from phenotypic plasticity (see Glossary 2.1) during development would, over many generations, lead to genetic change moderating this plasticity (Baldwin, 1896; Baldwin, 1902; Morgan, 1896; Osborn, 1897). Whether plasticity facilitates or slows down evolutionary diversification remains controversial (Pigliucci, 2006), but growing evidence suggests plasticity can influence the evolution of novel traits (Moczek et al., 2011; Pfennig et al., 2010). The potential role of phenotypic plasticity in shaping evolution was more comprehensively discussed a century later by West-Eberhard (2003), who emphasized that selection acts upon phenotypes, not genotypes. Phenotypes are not formed exclusively from genetic factors, but emerge from the integration of genetic, epigenetic (see Glossary 2.1) and environmental factors that act during development. While evolution is most commonly defined by changes in allele frequencies, a focus on genetic factors ignores the potential importance of environmental influences on phenotypic variation and evolution.

Other evolutionary biologists have also emphasized the potential importance of phenotypic plasticity in shaping evolution (Moczek et al., 2011; Pfennig et al., 2010; Pigliucci, 2006). Phenotypic plasticity can have large effects on fitness, allowing organisms to adapt to a changing environment and respond appropriately to inputs received during development. Plasticity itself may therefore be a target of selection (Nussey et al., 2005; Pigliucci, 2005; Van

Buskirk and Relyea, 1998), and phenotypic plasticity may also lead to the origin of novel phenotypes, preceding or even facilitating evolutionary change (Pfennig et al., 2010; Price et al., 2003). Plasticity-first evolution (see Glossary 2.1) (Levis and Pfennig, 2016) emphasizes the phenotype as the subject of selection, which can provide clarity for evolutionary models because phenotypic variation has a clear connection to natural selection.

The process of an environmentally induced phenotype leading to allele frequency change is known as genetic accommodation (see Glossary 2.1). Genetic accommodation can lead to either increased plasticity (such as the emergence of polyphenisms; see Glossary 2.1) or the fixation of an initially plastic trait, a special case known as genetic assimilation (see Glossary 2.1). While genetic accommodation has gained much theoretical support (Moczek et al., 2011; Pfennig et al., 2010; Pigliucci, 2006; West-Eberhard, 2003), it is difficult to test directly, which has likely influenced debate over the importance of plasticity-first evolution.

Arguments against plasticity-first evolution cite a lack of evidence for the molecular mechanisms enabling environmentally induced traits to become heritable in comparison to those that explain mutation-first evolution (Wray et al., 2014). Indeed, few examples of genetic accommodation have been elucidated to this level in the laboratory or in natural populations (e.g. Casasa and Moczek, 2018; Dworkin, 2005; Jones et al., 2017; Suzuki and Nijhout, 2006; reviewed in Renn and Schumer, 2013; Schlichting and Wund, 2014). However, it is unclear whether the current low number of genetic accommodation examples reflects actual rarity in nature, or a combination of low research effort together with a lack of known mechanisms. As our ability to study mechanisms of evolution in natural populations continues to improve, we expect more cases of genetic accommodation to be reported.

For many years, the molecular mechanisms that might enable environmentally induced traits to become heritable were unknown, and our lack of knowledge on the relationship between plasticity and genetic changes limited support for plasticity-first models. Recently, evidence for an interplay between plasticity-first mechanisms and elaboration of phenotypes via new mutations has been reported (Levis et al., 2018), highlighting the importance of considering both plasticity-first and mutation-first models in studies of evolutionary novelty. In addition, recent reports of epigenetic transgenerational inheritance (see Glossary 2.1) now provide plausible mechanisms through which genetic accommodation may act, priming the field to further investigate the role of plasticity-first mechanisms, including genetic accommodation, in

evolution.

In this Commentary, we briefly review some of these reports, describe the features of social insects that make them good models for studying genetic accommodation, and review current evidence consistent with a role for genetic accommodation in the evolution of eusociality (see Glossary 2.1). We end with a description of an empirical method to leverage the inherent plasticity of social insects to further study the mechanisms underlying eusocial evolution; we hope that this approach will lead to novel insights into the role that genetic accommodation has played in the evolution of social behavior.

Epigenetic Transgenerational Inheritance

There is new evidence for connections between the environment and adaptive phenotypic change across generations, as advances in the field of epigenetics provide plausible mechanisms for transgenerational inheritance. Many studies have demonstrated intergenerational (parent to offspring; see Glossary 2.1) or longer-lasting transgenerational effects, across plants, insects and mammals (Agrawal et al., 1999; Benito et al., 2018; Champagne, 2008; Dell and Rose, 1987; Gluckman et al., 2007; Ruden and Lu, 2008; Valtonen et al., 2012). While there is currently more evidence for intergenerational effects, other studies report evidence for transgenerational inheritance (Klosin et al., 2017; Siklenka et al., 2015). Most studies do not address the mechanisms of this inheritance (see Box 1), but a few have produced provocative associations with epigenetic changes such as DNA methylation (Dias and Ressler, 2014; Wei et al., 2014). For example, Dias and Ressler (2014) demonstrated that after adult male mice are subjected to odor fear conditioning, their offspring also exhibit fear of the same odor, despite no direct experience with the learning paradigm or odor. Additionally, they reported that both generations showed differences in DNA methylation at the locus encoding the olfactory receptor responsive to this odor, providing a putative mechanism of inheritance (Dias and Ressler, 2014).

Additional mechanisms of transgenerational inheritance have been identified, including small non-coding RNAs and chromatin remodeling (e.g. Gapp et al., 2014; Greer et al., 2011; reviewed in Houri-Zeevi and Rechavi, 2017; Jablonka and Raz, 2009). For example, male mice that engage in higher levels of voluntary wheel running show altered levels of microRNAs (miRNAs) and tRNA-derived RNAs in their sperm (Short et al., 2017). These males produce male offspring with reduced anxiety and suppressed juvenile fear memory, potentially mediated

through post-transcriptional gene regulation by the altered small RNAs in sperm (Short et al., 2017). Rodgers et al. (2015) demonstrated a direct effect of paternal miRNAs on offspring phenotypes through zygotic injection of nine paternal stress-related miRNAs, which led to reduced mRNA stores in zygotes and ultimately stress dysregulation phenotypes in offspring. Benito et al. (2018) also demonstrated a role for miRNAs in mediating synaptic plasticity in the offspring of male mice exposed to an environmental enrichment paradigm. Together, these studies demonstrate the possibly pervasive role of epigenetic mechanisms as mediators of transgenerational inheritance of environmentally induced phenotypes.

In addition to mediating environmental effects on phenotypes, epigenetic changes can also have direct effects on allele frequencies. For example, methylated cytosines make up nearly one-third of all germline and somatic point mutations as a result of increased rates of hydrolytic deamination at methylated cytosines when compared with unmethylated cytosines (Duncan and Miller, 1980; Shen et al., 1994). This leads to a depletion of CpG dinucleotides in genomes that undergo DNA methylation (Flores and Amdam, 2011). Environmental induction of methylation may thus lead to mutation of phenotypically relevant sites (Flores et al., 2013), altering the DNA sequence directly. Additionally, epigenetic marks for open chromatin lead to increased rates of transposable element insertion and meiotic recombination in maize (Liu et al., 2009). Thus, histone modifications that lead to accessible chromatin also increase the probability of genetic change. These mechanisms provide plausible links between phenotypic plasticity and mutation, demonstrating that these processes co-exist. While we are still in the early stages of understanding mechanisms by which parental experience shapes offspring phenotypes, mounting evidence suggests that mechanisms of transgenerational inheritance may be powerful modulators of phenotypic plasticity and, thus, evolution itself.

Box 1: Mechanisms of epigenetic inheritance

The environment experienced by parents can influence offspring phenotypes, either through direct exposure to an event (e.g. as germ cells or in utero) or as a result of altered parental care or other parent-mediated behaviors. When transmission is limited from parent to offspring (intergenerational epigenetic inheritance), epigenetic mechanisms are not required to explain inherited phenotypes, and work must be done to differentiate direct exposure to environmental stimuli from epigenetic inheritance. Transgenerational epigenetic inheritance, in

contrast, reflects long-lasting epigenetic effects in the absence of direct exposure to the stimulating environment. In recent years, concerns about the evolutionary relevance of epigenetic inheritance have been raised (Charlesworth et al., 2017), particularly in systems where causal connections have not been identified between epigenetic changes and the phenotype of interest. Still, many reports have identified epigenetic inheritance, and we briefly describe the main classes of mechanisms below (reviewed in Heard and Martienssen, 2014; Jablonka and Raz, 2009).

Chromatin-based mechanisms

Changes in chromatin, such as DNA methylation or histone modifications, are the best studied of all epigenetic inheritance mechanisms. In many cases, changes can be directly linked to differences in gene expression, and many laboratory assays (such as bisulfite sequencing and chromatin immunoprecipitation with sequencing) exist to readily measure chromatinbased epigenetic changes.

RNA-based mechanisms

Many types of RNA, including long non-coding RNA, small interfering RNA and microRNA, can persist across cell divisions and generations, altering DNA and histone modifications and/or directly affecting transcriptional and translational activity. Many of these RNA types have been found in germ line tissue, and manipulation of parentally mediated RNA can affect offspring phenotypes.

Self-sustaining regulatory loops

In bacteria and fungi, stable phenotypic states can involve transcriptional or post-transcriptional metabolic circuits that persist across generations. This was first reported in the lac operon of *Escherichia coli* (Novick and Weiner, 1957) and later demonstrated in other taxa.

Structural templates

Proteins, such as prions, which self-propagate by altering the structure of similar proteins, can transmit across cell divisions and have been shown to have transgenerational phenotypic effects in fungi. Protein chaperones may also mediate epigenetic variation by affecting protein folding across generations.

Phenotypic plasticity in social insects

Social insects have long been models for studying the role of environment on phenotype. Across species, a range of social forms is observed, from solitary to communal to complex eusocial species (Michener, 1974), with multiple independent origins of social phenotypes (Bourke, 2011). The breadth of behavioral plasticity across species provides unique opportunities to compare mechanisms of behavioral plasticity in a phylogenetic context. Additionally, many eusocial species exhibit extreme levels of plasticity between social castes, such as between individuals of different ages or between queens and workers.

Through the study of highly eusocial species such as the western honey bee (*Apis mellifera*), we know that environmental differences during development (e.g. larval nutrition) lead to caste differences, mediated by epigenetic mechanisms (Foret et al., 2012; Kucharski et al., 2008). Social insects have also played a critical role in uncovering the molecular basis of behavioral plasticity, with early transcriptomic studies of honey bees demonstrating for the first time that brain gene expression is predictive of behavioral state (Whitfield et al., 2003). Since then, many gene expression studies of social insects have identified transcriptomic differences associated with numerous phenotypic differences, including differences between queens and workers (Barchuk et al., 2007; Feldmeyer et al., 2014; Pereboom et al., 2005; Toth et al., 2007), differences between worker subcastes (Scharf et al., 2003; Whitfield et al., 2006) and different responses to socially relevant stimuli (Grozinger et al., 2003; Shpigler et al., 2017).

Brain transcriptional plasticity has been further modeled in the honey bee by using a large set of behavioral transcriptomic studies and reconstructing a brain transcriptional regulatory network. This network demonstrated context-dependent plasticity in the relationships between transcription factors and their target genes (Ament et al., 2012; Chandrasekaran et al., 2011), which is likely mediated through epigenetic mechanisms. Changes in DNA methylation and histone modifications have also been implicated in caste-related social behaviors in bees and ants (Herb et al., 2012, 2018; Lyko et al., 2010; Simola et al., 2015). Additionally, in a comparative study across 10 bee species, capturing multiple origins and elaborations of social behavior, sociality was correlated with increases in the occurrence of transcription factor binding sites and numbers of methylated genes, suggesting that eusocial lineages have an increased capacity for regulatory complexity (Kapheim et al., 2015b).

Although we have a good understanding of the mechanisms underlying behavioral

plasticity in eusocial insects, less is known about whether phenotypic plasticity is a precursor to eusocial evolution. West-Eberhard (2003) suggested that the worker caste, a defining feature of eusocial colonies (Wilson, 1971), evolved through environmental induction of phenotypically plastic traits. While much indirect evidence suggests an environmentally induced worker origin, including existing plasticity and experimental inducibility in related species (West-Eberhard, 2003), no direct evidence for a plasticity-first origin of worker castes has been demonstrated.

Levis and Pfennig (2016) outlined an empirical approach for assessing plasticity-first evolution in natural populations, with relevance for social insects. They describe characteristics of study systems well suited for studying genetic accommodation and plasticity-first evolution, including knowledge of phylogenetic relationships in the broader taxonomic group and five other criteria. Table 1 outlines these criteria and gives examples of how social insects match many of these characteristics. We review these characteristics in the next section, and describe how features of social insects make them amenable to studies of genetic accommodation.

Social insects as models for studying genetic accommodation

The multiple evolutionary origins of eusociality allow for comparisons of derived lineages with ancestral-proxy lineages (see Glossary 2.1) to resolve signals of eusocial evolution. A comparative approach has already been useful in finding patterns of evolution associated with social behavior, leveraging the natural variation in social forms present across bees (Kapheim et al., 2015b; Woodard et al., 2011). Across the hymenopteran social insects, eusociality has evolved at least 10 times (Bourke, 2011), with divergence times from solitary ancestors ranging from 20 to 100 million years ago (mya). While the selective forces acting on different lineages are not always known, some ecological circumstances such as nest-site limitation and parasitism have been implicated as determinants of group living (Gunnels et al., 2008; Langer et al., 2004), suggesting that these factors may be selective agents favoring eusocial evolution.

Social insects also exhibit quantifiable phenotypes that are amenable to molecular analysis, allowing us to explore how plasticity is achieved mechanistically. For example, a key component of eusociality is the presence of a reproductive division of labor between queens and workers. The degree of division of labor can be quantified using the skew in reproduction across adults (Sherman et al., 1995), with more complex eusocial species exhibiting greater skew. Reproductive skew can be measured in a laboratory setting, and has been used to assess whether

a division of labor occurs when typically solitary or subsocial (see Glossary 2.1) individuals are forced to cohabit (e.g. during forced association studies; see Glossary 2.1). Behavioral traits that are not related to reproduction can also be measured in a laboratory setting; such behaviors include excavation (Fewell and Page, 1999), foraging (Tenczar et al., 2014), guarding and nursing (Rittschof et al., 2014; Shpigler and Robinson, 2015). These quantifiable phenotypes could be useful in assessing the mechanisms underlying behavioral plasticity in eusocial colonies, and how these mechanisms may have changed throughout eusocial evolution.

Two social insect groups best match the characteristics outlined in Table 2.1 – bees and wasps. The multiple evolutionary origins of social behavior within each of these groups allows for phylogenetic comparisons of the mechanisms underlying eusocial evolution. In addition to multiple origins, closely related species display a range of social phenotypes, and growing genomic resources in these groups enable molecular studies of phenotypic plasticity (Ferreira et al., 2013; Jones et al., 2015; Kapheim et al., 2015b; Kocher et al., 2013; Standage et al., 2016). Already, these groups have shown promise in helping us to understand the potential role of genetic accommodation in eusocial evolution, as discussed in the next section.

Evidence for genetic accommodation in eusocial evolution

In addition to suggesting characteristics of ideal study systems for assessing plasticity-first evolution, Levis and Pfennig (2016) summarized four criteria necessary to establish that plasticity-first evolution has occurred, as listed below and in Table 2.2. Many studies of social insects demonstrate aspects of these criteria, although they were rarely formalized as studies of genetic accommodation (see references in Table 2.2). Below, we describe the criteria and provide examples from social insect studies that are consistent with each one. For social insects, the 'focal trait' (i.e. the phenotype under examination for testing predictions of genetic accommodation) we highlight is the presence of a reproductive division of labor between reproductive queen(s) and non-reproductive worker(s), as this is a defining feature of eusociality that is absent in non-eusocial ancestors.

Criterion 1: the focal trait can be environmentally induced in ancestral-proxy lineages

Reproductive division of labor has been induced experimentally in multiple solitary and subsocial species through forced association studies, where typically non-associating females are forced to cohabit. Many examples come from the small carpenter bees (Sakagami and Maeta,

1984, 1989, 1987) and sweat bees (Jeanson et al., 2005, 2008), groups that show high levels of social plasticity and may be especially useful for assessing genetic accommodation (Jones et al., 2017; Kocher and Paxton, 2014; Shell and Rehan, 2017). In some cases, a single species displays both solitary and social forms (Davison and Field, 2016; Smith et al., 2003; Soucy and Danforth, 2002), and in situations where these forms are environmentally determined, exploration of this variation may be particularly useful in studies of genetic accommodation.

Criterion 2: cryptic genetic variation is uncovered when ancestral-proxy lineages experience the derived environment

The capacity for variation in eusocial behavior can be uncovered through experimental manipulations of the social environment, as mentioned above for criterion 1. In the primarily solitary bee *Ceratina japonica*, artificial induction of multi-female nests results in a division of labor among females, including skew in reproduction between adults (Sakagami and Maeta, 1984). Similar multi-female nest induction in the related bee *Ceratina flavipes* results in fewer nests that successfully rear brood, but those that do also show evidence of a rudimentary caste system (Sakagami and Maeta, 1987). Whether successful multi-female nests contain females with cryptic genetic variation (see Glossary 2.1) enabling division of labor has not been explored. However, populations of the facultatively eusocial *Lasioglossum albipes* with different social forms display genetic differentiation, suggesting that variation in sociality may be facilitated by genetic variation (Kocher et al., 2013). As genomic tools become available for many social insect species, investigations of how cryptic genetic variation influences environmentally induced trait variation will be a critical step in studies of genetic accommodation in social insects.

Criterion 3: the focal trait exhibits evidence of evolutionary change in regulation and/or form in derived lineages

Levis and Pfennig (2016) suggest that both genetically accommodated and assimilated traits will exhibit changes in the slope of the reaction norm (see Glossary 2.1) when comparing derived with ancestral-proxy lineages. Genetically assimilated traits would additionally have fixed reaction norms across different environments, compared with more flexible phenotypes in the ancestral-proxy species (Levis and Pfennig, 2016). Finally, the mechanisms underlying changes in reaction norms should be evident, such as changes in hormonal signaling, cis-regulatory elements and alternative splicing (Levis and Pfennig, 2016). This criterion is well

supported from multiple indirect lines of evidence in social insects. Among obligately eusocial insects, division of labor is fixed compared with ancestral-proxy lineages in which colonies can exist in multiple states, including those without division of labor (e.g. the colony-founding phase of single *Bombus* queens) or species with both solitary and social forms (e.g. *Megalopta genalis* and *Lasioglossum albipes*: Kocher and Paxton, 2014). Many caste-related genes show evidence of positive selection in ants (*Solenopsis* spp.: Hunt et al., 2012; *Temnothorax longispinosus*: Feldmeyer et al., 2014) and bees (*Apis mellifera*: Harpur et al., 2014; Hunt et al., 2010; *Megalopta genalis*: Jones et al., 2017; cross-species comparison: Woodard et al., 2011) relative to genes not related to caste expression. These results suggest a change in usage of these genes relative to their expression in solitary ancestors that affects their evolutionary rates. In addition, computational analysis shows that changes in gene regulatory capacity correlate with the level of eusociality, in two different contexts. First, there are predicted increases in the strength and prevalence of transcription factor binding sites in gene promoters of species with increased levels of eusociality (Kapheim et al., 2015b). Second, there are greater predicted numbers of methylated genes associated with increased levels of eusociality, and the role of gene methylation in controlling expression and splicing in social insects further suggests greater regulatory tuning in social species (Kapheim et al., 2015b). These cis-regulatory and transcription factor differences suggest that genes related to eusociality have experienced evolutionary change in both sequence and regulation.

Criterion 4: the focal trait exhibits evidence of having undergone adaptive refinement in derived lineages

If selection has acted to increase the frequency with which a trait is expressed through genetic accommodation, that trait should experience adaptive refinement (see Glossary 2.1) as a result of more consistent exposure to selection (West-Eberhard, 2003). Therefore, genetically accommodated traits in derived lineages should be superior versions of the trait compared with those in ancestral-proxy lineages (Levis and Pfennig, 2016). In the majority of social insect species, a comparison of fitness in solitary and eusocial nests is not possible; however, alternative social strategies of the facultatively eusocial *Megalopta genalis* were found to have similar levels of fitness and to co-exist in evolutionary models based on field-based parameters, potentially explaining the maintenance of facultative eusociality in this species (Kapheim et al., 2015a). Further studies that rear solitary and eusocial species in competition may help elucidate

whether eusocial traits have undergone adaptive refinement more directly. In addition, several indirect lines of evidence support the idea that eusociality has undergone adaptive refinement in derived lineages, including extensive elaborations of form and function in queen and workers castes. In some derived lineages, workers have lost the ability to mate and are sterile, while queens have reproductive capacities that are orders of magnitude greater than those of their ancestral-proxy counterparts. This specialization of castes would likely not be possible without the fixed sociality present in these complex eusocial species, where queen and worker traits have been exposed to selection for millions of years. Species with flexible castes or facultative eusociality are less consistently exposed to selection, and do not display the same features as complex eusocial species, suggesting a greater capacity for adaptive refinement in lineages with complex and obligate eusociality. Elaborate chemical communication systems have also evolved in many of these complex eusocial groups, leading to less overt competition over reproduction and highly specialized nestmate recognition systems, providing further evidence for adaptive refinement in derived lineages.

Eusocial engineering

Most studies providing empirical support for genetic accommodation employ artificial selection on experimentally induced phenotypes, some of which may not occur in nature (Suzuki and Nijhout, 2006; Waddington, 1942, 1953). Other empirical studies demonstrate phylogenetic relationships between environmentally sensitive phenotypes in ancestral lineages and more fixed phenotypes in derived lineages (Heil et al., 2004; Santana and Dumont, 2009; West-Eberhard, 2003), but do not examine the underlying mechanisms or show evidence of allele frequency change associated with fixation of the phenotype. Here, we outline a new method that builds upon and extends these approaches.

As discussed above, evidence across the social insects suggests that phenotypic plasticity may have facilitated the origin of eusociality through genetic accommodation. However, many of these lines of evidence come from different lineages and distinct origins of eusocial behavior, making it difficult to assess whether plasticity-first evolution has taken place. We suggest that particular social insect groups can be used to test the predictions of plasticity-first evolution, coupling manipulative experiments with deep molecular probing of the mechanisms involved in shifts between social forms. Specifically, we propose using experimental environmental

induction of novel social traits (as in Table 2.2, criterion 1) to test the hypothesis that genetic accommodation can act on existing plasticity in social evolution. We call this approach 'eusocial engineering' (see Glossary 2.1), and hope that it will enable better understanding of how environmentally induced phenotypes may be involved in gains and losses of eusocial behavior.

Fig. 2.1 outlines two approaches to eusocial engineering. The forward eusocial engineering technique (Fig. 2.1A) is as follows. For a species closer to the solitary end of the social spectrum with potential for non-lethal association of females (i.e. communal or facultatively eusocial species), nests are established with pairs or trios of age-matched, mated females of the same generation (mimicking a semisocial origin of eusociality; Michener, 1974), or mother and daughter(s) groups (mimicking a subsocial origin; Michener, 1974). Nests are monitored until successful generation of brood, which may occur in a subset of nests. After emergence of offspring, founding females are analyzed to examine changes in gene expression and epigenetic regulation, and offspring are used to establish new nests to artificially select for high fitness in the induced social environment. While the molecular approaches we suggest may be expensive and require practical considerations, sequencing costs and more tractable techniques for non-model organisms are being developed with increased frequency, bringing these methods within reach of many biologists.

Throughout a study, observations would identify behavioral division of labor, and successful females would be preserved for physiological and molecular measurements. Genotyping of offspring (for females, this would be conducted after their own nests have completed a brood cycle) would identify which founding female successfully reproduced, with ovarian dissection of all females to assess reproductive potential. Transcriptomics on collected foundresses would look for differences in gene expression associated with variation in phenotype induced by the social environment. Genes with plastic expression associated with the social induction would be candidates for selection through genetic accommodation. In addition to differences in gene expression, novel regulation of genes can be assessed by examining differences in chromatin accessibility and methylation using a number of techniques (e.g. ATAC-seq: Buenrostro et al., 2013; ChIP-seq: Barski et al., 2007; bisulfite sequencing: Clark et al., 1994) to address which molecular mechanisms have enabled the plasticity in behavior induced by the novel social environment. Sampling of females throughout many generations of artificial selection would allow a direct test of genetic accommodation by assessing whether

candidate genes exhibit allele frequency change or novel variants following selection. In early generations, transcriptomic differences may be observed without reinforcement from epigenetic mechanisms. Epigenetic reinforcement may be expected later, with eventual accommodation of changes resulting in allele frequency change relative to the starting population of individuals. Careful genetic surveying of populations before and after the selection regime would be necessary to understand the role that cryptic genetic variation might play in the emergence of social traits. Longitudinal studies of laboratory bacterial evolution provide excellent experimental guides for this work (e.g. Bohannan and Lenski, 2000).

Reverse eusocial engineering (Fig. 2.1B) would use species with higher social complexity, and involve the initiation of nests with single females followed by selection on those individuals that successfully reproduce under solitary conditions. As with the forward eusocial engineering scheme above, molecular techniques would be used to assess mechanisms associated with reversions to solitary living, as well as test for genetic accommodation for the environmentally induced solitary phenotype. An additional approach could use selected lines from forward eusocial engineering as a starting point, with molecular monitoring to address how eusociality may be lost when females are forced to rear offspring in isolation.

The foundational behavioral aspects of the forward eusocial engineering approach have already been established, as forced association studies have been successful in multiple species of both ants and bees (Fewell and Page, 1999; Rissing and Pollock, 1986; Sakagami and Maeta, 1987). Reverse eusocial engineering has less behavioral precedence, but eusocial behavior has been lost multiple times in some groups (e.g. Danforth et al., 2003), suggesting that reversion to solitary behavior may be a common phenomenon and therefore important to study. Other work has successfully manipulated the social environment (Robinson et al., 1989, 1992; Ross and Keller, 2002), including changing colony demographics in ‘pseudomutant’ colonies and comparing the performance of this artificial construct with naturally formed colonies (Wilson, 1985), as well as many instances of queen removals to induce worker reproduction across ants, bees and wasps (e.g. Dietemann and Peeters, 2000; Jones et al., 2017; Reeve and Gamboa, 1987). What is unique about the eusocial engineering approach is the coupling of these manipulative studies with artificial selection and multiple genetic, transcriptomic and epigenetic monitoring approaches, enabling real-time tracking of the plastic and heritable components of environmentally induced traits. With recent advances in transgenic approaches in social insects

(Schulte et al., 2014; Trible et al., 2017; Yan et al., 2017), eusocial engineering could be followed by genome editing to directly test the effects of any discovered genetic and epigenetic variants that are found to be associated with transitions in social behavior. We expect that eusocial engineering will be valuable in testing the role of genetic accommodation and phenotypic plasticity in eusocial evolution.

Outside of social insects, phenotype engineering has been previously introduced and implemented in birds; hormonal manipulations were utilized to change behavior and physiology in the dark-eyed junco (Ketterson and Nolan, 1992). This work suggests an approach similar to eusocial engineering may also be fruitful in testing genetic accommodation in non-insects.

Conclusions

An emphasis on plasticity-first evolution does not diminish the importance of mutation-first mechanisms of evolutionary change, but rather adds a potential avenue for scientists to explore for comprehensive analyses of the evolution of complex traits. Selection pressure is agnostic to the mechanism leading to the phenotype, such that individuals that inherited fitness benefits via reversible epigenetic means would survive equally well as individuals with a genetic mutation permanently altering the phenotype, all else being equal. In changing environmental conditions, however, a plastic response may be more advantageous and lead to maintenance of a transgenerational mechanism of inheritance. By contrast, if a population experiences a stable environment, individuals with a fixed and genetically determined phenotype may ultimately edge out those with plastic responses, either because of the costs of plasticity or because of differences in the reliability of phenotypic expression (DeWitt et al., 1998). In this case, the population may experience allele frequency change, completing the plasticity-first model of evolution.

Plasticity-first models of evolution arose before our current knowledge of the possible mechanisms of genetic accommodation, but growing understanding of epigenetics and transgenerational plasticity allows us to now test the predictions of these models. Social insects are well suited for this goal, and an empirical approach that combines behavioral manipulations with ‘omics work will open the door to understanding how transcriptional plasticity in the ancestors of eusocial species may have facilitated the evolution of eusocial traits. With this approach applied more broadly in other taxa, we believe that the time has come to consider the role of genetic accommodation more rigorously, in order to determine its significance as a driver

of evolutionary change. In doing so, we will not only broaden our understanding of the role of phenotypic plasticity in the origin and elaboration of novel traits but also provide a framework by which multiple modes of evolution may work in concert to influence adaptation.

Acknowledgments

I would like to thank members of the Robinson lab, M. B. Sokolowski, M. J. West-Eberhard and several anonymous reviewers for comments that improved this text. This work was supported by National Science Foundation IGERT (PI Suarez, UIUC) and GRFP fellowships, as well as National Science Foundation and National Institutes of Health grants to my co-author Gene Robinson.

Glossary, Figure, and Tables

Glossary 2.1

Adaptive refinement: Increased fitness relative to an ancestral state, possibly facilitated by constitutive expression in the derived lineage related to the ancestral condition.

Ancestral-proxy lineages: Lineages closely related to the derived lineage of interest that display the ancestral character state for the trait of interest; these lineages lack the derived trait of interest while in their natural (ancestral) environment, but may exhibit plasticity for the trait when exposed to a novel environment (indicating pre-existing plasticity).

Cryptic genetic variation: Genetic variation in a population that does not currently contribute to phenotypic outcomes, but that may modify phenotypes following environmental change or new epistatic interactions with novel alleles.

Epigenetic: Referring to any feature of chromatin, DNA or other cellular features other than the DNA sequence itself that may influence gene expression and function, and may lead to heritable changes in transcriptional activity across cell divisions and/or generations.

Eusocial engineering: A proposed coupling of forced association studies with transcriptomics, epigenomics and other molecular analyses to test the mechanisms of genetic accommodation that may have acted on ancestral plasticity in social evolution.

Eusociality: Defined by (1) a reproductive division of labor (a queen that reproduces, workers that do not), (2) overlapping adult generations (often a mother and her daughters or sisters and their offspring), and (3) cooperative care of brood.

Forced association study: A study involving experimentally induced group formation, including of naturally solitary individuals, members of different natural groups or artificially created age classes of individuals; used to study emergent properties of groups or to remove confounds of developmental experience in studying social behavior.

Genetic accommodation: A process by which initially environmentally induced and plastic phenotypes are selected upon, resulting in heritable variation influencing the expression of those phenotypes; genetic accommodation can lead to increased plasticity for the trait (including the emergence of polyphenisms) or decreased plasticity (see *genetic assimilation*).

(Glossary 2.1 continued)

Genetic assimilation: A special case of genetic accommodation, where initially plastic traits become fixed through selection on one or more alternative genotypes; this results in reduced phenotypic plasticity over evolutionary time.

Intergenerational inheritance: Transference of environmentally mediated epigenetic changes from parent to offspring.

Mutation-first evolution: A mechanism of evolution in which a novel mutation or novel allele in the population alters a phenotype under selection, leading to changes in allele frequencies; contrast with plasticity-first evolution.

Phenotypic plasticity: The ability of a single genotype to produce multiple phenotypes in response to epigenetic or environmental conditions.

Plasticity-first evolution: A mechanism of evolution in which novel, environmentally sensitive phenotypic variation (i.e. phenotypic plasticity) provides the initial substrate for selection, followed by changes in allele frequencies in the population through selection on cryptic genetic variation (see above) underlying the phenotypic plasticity and/or accommodation on the newly selected trait; contrast with mutation-first evolution.

Polyphenism: Discrete phenotypic forms arising from phenotypic plasticity; often phenotypic forms are very distinct, such as queen and worker castes of complex eusocial insects or color morphs of some butterflies.

Reaction norm: The pattern of expressed phenotypes for a given genotype across one or more environmental variables, typically represented graphically; an individual shows phenotypic plasticity for a trait if the slope of the reaction norm is non-zero, indicating an interaction effect between genotype and environment for the phenotype of interest.

Subsocial: Adult females that protect and/or feed their developing offspring, but disperse or die prior to offspring emergence such that there is no adult generational overlap and no division of labor among adults.

Transgenerational inheritance: Transference of environmentally mediated epigenetic changes across more than two generations.

Figure 2.1

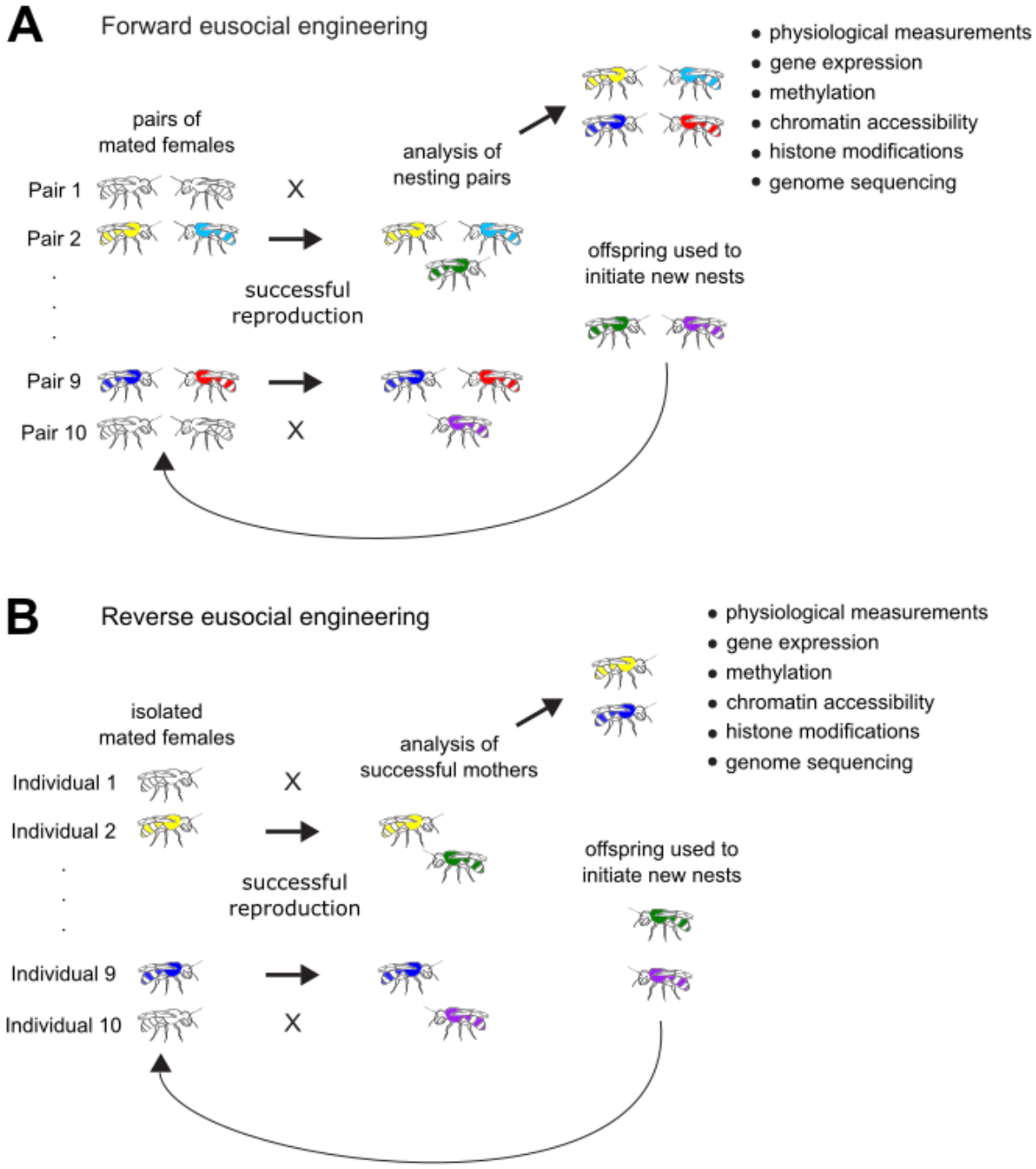


Figure 2.1. Schematic diagram outlining forward and reverse eusocial engineering. (A) Forward eusocial engineering. The first generation includes multiple pairs of individuals, some of which may successfully produce offspring (colored pairs, arrows; some pairs fail to produce

(Figure 2.1 caption continued)

offspring, denoted by ‘X’). Offspring of these successful pairs (shown in green and purple) will be paired again in the next generation, to either successfully produce another generation or fail to produce offspring. This pairing and selection scheme continues for many generations, with sampling of successful pairs throughout to identify transcriptomic, epigenomic and allele frequency changes associated with selection for cooperation. (B) Reverse eusocial engineering. The first generation includes isolated females, some of which may successfully produce offspring (shown in yellow and blue, arrows; some individuals fail to produce offspring in isolation, denoted by ‘X’). Offspring of successful females (shown in green and purple) will be isolated again for the next generation. Much like in forward eusocial engineering, successful individuals can be sampled throughout to identify transcriptomic, epigenomic and allele frequency changes associated with selection for solitary reproduction. Many variations of this scheme are possible depending upon the species of interest.

Table 2.1. Characteristics of study systems well suited to studies of plasticity-first evolution (adapted from Levis and Pfennig, 2016), as well as select examples from social insects.

Characteristic	Examples from social insects	Taxa/references
Knowledge of phylogenetic relationships in the broader taxonomic group	Well-resolved phylogenies for many groups of social insects	(Branstetter et al., 2017; Johnson et al., 2013; Romiguier et al., 2016)
Multiple parallel derived lineages, with variable divergence times	Bees: 4–6 independent origins, ranging from 20 to 65 mya Wasps: 4 independent origins, unknown origin dates for all but one lineage (100 mya)	(Bourke, 2011; Brady et al., 2006; Cameron and Mardulyn, 2001; Chenoweth et al., 2007; Schwarz et al., 2007; Thompson and Oldroyd, 2004)
Knowledge of ecological circumstances and selective agents acting on lineages	Nest-site limitation: habitat saturation selects for non-dispersal and group living Parasitism/predation: groups are better protected from parasites or predators	*Exoneura nigrescens* (Langer et al., 2004), *Mischocyttarus mexicanus* (Gunnels et al., 2008), *Xylocopa sulcatipes* (Stark, 1992), *Megalopta genalis* (Smith et al., 2003)
Quantifiable trait that can be induced under laboratory conditions	Reproductive skew, behavioral castes, division of labor; induced in forced association studies or environmental manipulations, or observed naturally in observation nests of social species	*Veromessor pergandei* (Rissing and Pollock, 1986), *Ceratina flavipes* (Sakagami and Maeta, 1987), *Pogonomyrmex barbatus* (Fewell and Page, 1999), *Apis mellifera* (Robinson et al., 1989), *Megalopta genalis* (Jones et al., 2017)
Adequate genomic resources to investigate molecular underpinnings	Genomes per group (from NCBI, accessed 25 February 2018): Ants: 18 Bees: 15 Vespid wasps: 2 Termites: 3	*Acromyrmex echinatior, Atta cephalotes, Atta colombica, Camponotus floridanus, Cyphomyrmex costatus, Dinoponera quadriceps, Harpegnathos saltator, Lasius niger, Monomorium pharaonis, Oooceraea biroi, Pogonomyrmex barbatus, Pseudomyrmex gracilis, Solenopsis invicta, Trachymyrmex cornetzi, Trachymyrmex septentrionalis, Trachymyrmex zeteki, Vollenhovia emeryi, Wasmannia auropunctata;*

(Table 2.1 continued)

		Apis cerana, Apis dorsata, Apis florea, Apis mellifera, Bombus impatiens, Bombus terrestris, Ceratina calcarata, Dufourea novaeangliae, Eufriesea mexicana, Euglossa dilemma, Habropoda laboriosa, Lasioglossum albipes, Lepidotrigona ventralis, Megachile rotundata, Melipona quadrifasciata; Polistes canadensis, Polistes dominula; Cryptotermes brevis, Nasutitermes exitiosus, Zootermopsis nevadensis
Features amenable to lab rearing, including fast generation time, numerous offspring, etc.	High variability across groups, but many species are easy to maintain in lab colonies, and social species produce many offspring	*Solenopsis invicta* (Banks et al., 1981), *Bombus impatiens* (Cnaani et al., 2002), *Polistes metricus* (Daugherty et al., 2011), *Polistes fuscatus* (Gibo, 1974)

Table 2.2. Criteria for establishing plasticity-first evolution in natural populations (adapted from Levis and Pfennig, 2016), with examples from social insects.

Criterion	Example from social insect literature	Taxa/references
Focal trait can be environmentally induced in ancestral-proxy lineages	Induction of castes in artificial multi-female nests of solitary or subsocial species	*Ceratina japonica* (Sakagami and Maeta, 1984, 1987), *Ceratina okinawana* (Sakagami and Maeta, 1989), *Lasioglossum* spp. (Jeanson et al., 2005, 2008)
Cryptic genetic variation is uncovered when ancestral-proxy lineages experience the derived environment	Indirect evidence: inducible social phenotypes, and some genetic differentiation between social and solitary forms	*Ceratina japonica* (Sakagami and Maeta, 1984); *Ceratina flavipes* (Sakagami and Maeta, 1987); *Lasioglossum albipes* (Kocher et al., 2013)
Focal trait exhibits evidence of evolutionary change in regulation/form in derived lineages	Evidence of positive selection on caste-related genes in multiple lineages Social lineages exhibit increases in transcription factor binding site strength/presence and increased numbers of methylated genes	*Solenopsis* spp. (Hunt et al., 2012), *Temnothorax longispinosus* (Feldmeyer et al., 2014), *Apis mellifera* (Harpur et al., 2014; Hunt et al., 2010), *Megalopta genalis* (Jones et al., 2017), 10 bee species comparison (Kapheim et al., 2015b)
Focal trait exhibits evidence of adaptive refinement in derived lineages	Highly specialized queen and worker castes in complex eusocial lineages	Most ant species, *Apis mellifera*

References

Agrawal, A. A., Laforsch, C. and Tollrian, R. (1999). Transgenerational induction of defences in animals and plants. *Nature* 401, 60–63.

Ament, S., Wang, Y., Chen, C.-C., Blatti, C., Hong, F., Liang, Z., Negre, N., White, K., Rodriquez-Zas, S., Mizzen, C., et al. (2012). The transcription factor *Ultraspiracle* influences honey bee social behavior and behavior-related gene expression. *PLoS Genet.* 8, e1002596.

Baldwin, J. M. (1896). A new factor in evolution. *Am. Nat.* 30, 441–451.

Baldwin, J. M. (1902). *Development and Evolution*. London: MacMillan Co.

Banks, W. A., Lofgren, C. S., Jouvenaz, D. P., Stringer, C. E., Bishop, P. M., Williams, D. F., Wojcik, D. P. and Glancey, B. M. (1981). Techniques for collecting, rearing, and handling imported fire ants. *U.S. Dep. Agric. Sci. Educ. Adm* 21, 1–9.

Barchuk, A. R., Cristino, A. S., Kucharski, R., Costa, L. F., Simões, Z. L. P. and Maleszka, R. (2007). Molecular determinants of caste differentiation in the highly eusocial honeybee *Apis mellifera*. *BMC Dev. Biol.* 7, 70.

Barski, A., Cuddapah, S., Cui, K., Roh, T. Y. and Schones, D. E. (2007). High-resolution profiling of histone methylations in the human genome. *Cell* 129, 823–37.

Benito, E., Kerimoglu, C., Ramachandran, B., Pena-Centeno, T., Jain, G., Stilling, R. M., Islam, M. R., Capece, V., Zhou, Q., Edbauer, D., et al. (2018). RNA-dependent intergenerational inheritance of enhanced synaptic plasticity after environmental enrichment. *Cell Rep.* 23, 546–554.

Bohannan, B. J. M. and Lenski, R. E. (2000). Linking genetic change to community evolution: Insights from studies of bacteria and bacteriophage. *Ecol. Lett.* 3, 362–377.

Bourke, A. F. G. (2011). *Principles of Social Evolution*. Oxford: Oxford University Press.

Brady, S. G., Sipes, S., Pearson, A. and Danforth, B. N. (2006). Recent and simultaneous origins of eusociality in halictid bees. *Proc. R. Soc. B Biol. Sci.* 273, 1643–1649.

Branstetter, M. G., Danforth, B. N., Pitts, J. P., Faircloth, B. C., Ward, P. S., Buffington, M. L., Gates, M. W., Kula, R. R. and Brady, S. G. (2017). Phylogenomic insights into the evolution of stinging wasps and the origins of ants and bees. *Curr. Biol.* 27, 1019–1025.

Buenrostro, J., Giresi, P., Zaba, L., Chang, H. and Greenleaf, W. (2013). Transposition of native chromatin for fast and sensitive epigenomic profiling of open chromatin, DNA-binding proteins and nucleosome position. *Nat. Methods* 10, 1213–1218.

Cameron, S. A. and Mardulyn, P. (2001). Multiple molecular data sets suggest independant origins of highly eusocial behavior in bees (Hymenoptera: Apinae). *Syst. Biol.* 50, 194–214.

Carroll, S. B. (2008). Evo-Devo and an expanding evolutionary synthesis: A genetic theory of morphological evolution. *Cell* 134, 25–36.

Casasa, S. and Moczek, A. P. (2018). The role of ancestral phenotypic plasticity in evolutionary diversification: population density effects in horned beetles. *Anim. Behav.* 137, 53–61.

Champagne, F. A. (2008). Epigenetic mechanisms and the transgenerational effects of maternal care. *Front. Neuroendocrinol.* 29, 386–397.

Chandrasekaran, S., Ament, S. A., Eddy, J. A., Rodriguez-Zas, S. L., Schatz, B. R., Price, N. D. and Robinson, G. E. (2011). Behavior-specific changes in transcriptional modules lead to distinct and predictable neurogenomic states. *Proc. Natl. Acad. Sci.* 108, 18020.

Charlesworth, D., Barton, N. H. and Charlesworth, B. (2017). The sources of adaptive variation. *Proc. R. Soc. B Biol. Sci.* 284, 20162864.

Chenoweth, L. B., Tierney, S. M., Smith, J. A., Cooper, S. J. B. and Schwarz, M. P. (2007). Social complexity in bees is not sufficient to explain lack of reversions to solitary living over long time scales. *BMC Evol. Biol.* 7, 246.

Clark, S. J., Harrison, J., Paul, C. L. and Frommer, M. (1994). High sensitivity mapping of methylated cytosines. *Nucleic Acids Res.* 22, 2990–7.

Cnaani, J., Schmid-Hempel, R. and Schmidt, J. O. (2002). Colony development, larval development and worker reproduction in *Bombus impatiens* Cresson. *Insectes Soc.* 49, 164–170.

Danforth, B. N., Ji, S. and Conway, L. (2003). Phylogeny of eusocial *Lasioglossum* reveals multiple losses of eusociality within a primitively eusocial clade of bees (Hymenoptera: Halictidae). *Syst. Biol.* 52, 23–36.

Daugherty, T. H. F., Toth, A. L. and Robinson, G. E. (2011). Nutrition and division of labor: Effects on foraging and brain gene expression in the paper wasp *Polistes metricus*. *Mol. Ecol.* 20, 5337–5347.

Davison, P. J. and Field, J. (2016). Social polymorphism in the sweat bee *Lasioglossum (Evylaeus) calceatum*. *Insectes Soc.* 63, 327–338.

Dell, P. A. and Rose, F. D. (1987). Transfer of effects from environmentally enriched and impoverished female rats to future offspring. *Physiol. Behav.* 39, 187–190.

DeWitt, T. J., Sih, A. and Wilson, D. S. (1998). Costs and limits of phenotypic plasticity. *Trends Ecol. Evol.* 13, 77–81.

Dias, B. G. and Ressler, K. J. (2014). Parental olfactory experience influences behavior and neural structure in subsequent generations. *Nat. Neurosci.* 17, 89–96.

Dietemann, V. and Peeters, C. (2000). Queen influence on the shift from trophic to reproductive eggs laid by workers of the ponerine ant *Pachycondyla apicalis*. *Insectes Soc.* 47, 223–228.

Duncan, B. K. and Miller, J. H. (1980). Mutagenic deamination of cytosine residues in DNA. *Nature* 287, 560–561.

Dworkin, I. (2005). A study of canalization and developmental stability in the sternopleural bristle system of *Drosophila melanogaster*. *Evolution* 59, 1500–1509.

Feldmeyer, B., Elsner, D. and Foitzik, S. (2014). Gene expression patterns associated with caste and reproductive status in ants: Worker-specific genes are more derived than queen-specific ones. *Mol. Ecol.* 23, 151–161.

Ferreira, P. G., Patalano, S., Chauhan, R., Ffrench-Constant, R., Gabaldón, T., Guigó, R. and Sumner, S. (2013). Transcriptome analyses of primitively eusocial wasps reveal novel insights into the evolution of sociality and the origin of alternative phenotypes. *Genome Biol.* 14, R20.

Fewell, J. H. and Page R E, J. (1999). The emergence of division of labour in forced associations of normally solitary ant queens. *Evol. Ecol. Res.* 1, 537–548.

Flores, K. B. and Amdam, G. V. (2011). Deciphering a methylome: what can we read into patterns of DNA methylation? *J. Exp. Biol.* 214, 3155–3163.

Flores, K. B., Wolschin, F. and Amdam, G. V. (2013). The role of methylation of DNA in environmental adaptation. *Integr. Comp. Biol.* 532, 359–372.

Foret, S., Kucharski, R., Pellegrini, M., Feng, S., Jacobsen, S. E., Robinson, G. E. and Maleszka, R. (2012). DNA methylation dynamics, metabolic fluxes, gene splicing, and alternative phenotypes in honey bees. *Proc. Natl. Acad. Sci.* 109, 4968–4973.

Gapp, K., Jawaid, A., Sarkies, P., Bohacek, J., Pelczar, P., Prados, J., Farinelli, L., Miska, E. and Mansuy, I. M. (2014). Implication of sperm RNAs in transgenerational inheritance of the effects of early trauma in mice. *Nat. Neurosci.* 17, 667–669.

Gibo, D. L. (1974). A laboratory study on the selective advantage of foundress associations in *Polistes fuscatus* (Hymenoptera: Vespidae). *Can. Entomol.* 106, 101–106.

Gluckman, P. D., Hanson, M. A. and Beedle, A. S. (2007). Non-genomic transgenerational inheritance of disease risk. *BioEssays* 29, 145–154.

Greer, E. L., Maures, T. J., Ucar, D., Hauswirth, A. G., Mancini, E., Lim, J. P., Benayoun, B. A., Shi, Y. and Brunet, A. (2011). Transgenerational epigenetic inheritance of longevity in *Caenorhabditis elegans*. *Nature* 479, 365–371.

Grozinger, C. M., Whitfield, C. W., Robinson, G. E. and Sharabash, N. M. (2003). Pheromone-mediated gene expression in the honey bee brain. *Proc. Natl. Acad. Sci.* 100, 14519–14525.

Gunnels, C. W., Dubrovskiy, A. and Avalos, A. (2008). Social interactions as an ecological constraint in a eusocial insect. *Anim. Behav.* 75, 681–691.

Harpur, B. A., Kent, C. F., Molodtsova, D., Lebon, J. M. D., Alqarni, A. S., Owayss, A. A., and Zayed, A. (2014). Population genomics of the honey bee reveals strong signatures of positive selection on worker traits. *Proc. Natl. Acad. Sci.* 111, 2614–2619.

Heard, E. and Martienssen, R. A. (2014). Transgenerational epigenetic inheritance: Myths and mechanisms. *Cell* 157, 95–109.

Heil, M., Greiner, S., Meimberg, H., Krüger, R., Noyer, J.-L., Heubl, G., Eduard Linsenmair, K. and Boland, W. (2004). Evolutionary change from induced to constitutive expression of an indirect plant resistance. *Nature* 430, 205–208.

Herb, B. R., Wolschin, F., Hansen, K. D., Aryee, M. J., Langmead, B., Irizarry, R., Amdam, G. V and Feinberg, A. P. (2012). Reversible switching between epigenetic states in honeybee behavioral subcastes. *Nat. Neurosci.* 15, 1371.

Herb, B. R., Shook, M. S., Fields, C. J. and Robinson, G. E. (2018). Defense against territorial intrusion is associated with DNA methylation changes in the honey bee brain. *BMC Genomics* 19, 216.

Houri-Zeevi, L. and Rechavi, O. (2017). A matter of time: Small RNAs regulate the duration of epigenetic inheritance. *Trends Genet.* 33, 46–57.

Hunt, B. G., Wyder, S., Elango, N., Werren, J. H., Zdobnov, E. M., Yi, S. V and Goodisman, M. A. D. (2010). Sociality is linked to rates of protein evolution in a highly social insect. *Mol. Biol. Evol.* 27, 497–500.

Hunt, B. G., Ometto, L., Keller, L. and Goodisman, M. A. D. (2012). Evolution at two levels in fire ants: The relationship between patterns of gene expression and protein sequence evolution. *Mol. Biol. Evol.* 30, 263–271.

Jablonka, E. and Raz, G. (2009). Transgenerational epigenetic inheritance: prevalence, mechanisms, and implications for the study of heredity and evolution. *Q. Rev. Biol.* 84, 131–176.

Jeanson, R., Kukuk, P. F. and Fewell, J. H. (2005). Emergence of division of labour in halictine bees: Contributions of social interactions and behavioural variance. *Anim. Behav.* 70, 1183–1193.

Jeanson, R., Clark, R. M., Holbrook, C. T., Bertram, S. M., Fewell, J. H. and Kukuk, P. F. (2008). Division of labour and socially induced changes in response thresholds in associations of solitary halictine bees. *Anim. Behav.* 76, 593–602.

Johnson, B. R., Borowiec, M. L., Chiu, J. C., Lee, E. K., Atallah, J. and Ward, P. S. (2013). Phylogenomics resolves evolutionary relationships among ants, bees, and wasps. *Curr. Biol.* 23, 2058–2062.

Jones, B. M., Wcislo, W. T. and Robinson, G. E. (2015). Developmental transcriptome for a facultatively eusocial bee, *Megalopta genalis*. *G3* 5, 2127.

Jones, B. M., Kingwell, C. J., Wcislo, W. T. and Robinson, G. E. (2017). Caste-biased gene expression in a facultatively eusocial bee suggests a role for genetic accommodation in the evolution of eusociality. *Proc. R. Soc. B Biol. Sci.* 284, 20162228.

Kapheim, K. M., Nonacs, P., Smith, A. R., Wayne, R. K. and Wcislo, W. T. (2015a). Kinship, parental manipulation and evolutionary origins of eusociality. *Proc. R. Soc. B Biol. Sci.* 282, 20142886.

Kapheim, K. M., Pan, H., Li, C., Salzberg, S. L., Puiu, D., Magoc, T., Robertson, H.M., Hudson, M.E., Venkat, A., Fischman, B.J. et al. (2015b). Genomic signatures of evolutionary transitions from solitary to group living. *Science* 348, 1139–1143.

Ketterson, E. D. and Nolan, V. (1992). Hormones and life histories: An integrative approach. *Am. Nat.* 140, S33–S62.

Klosin, A., Casas, E., Hidalgo-Carcedo, C., Vavouri, T. and Lehner, B. (2017). Transgenerational transmission of environmental information in *C. elegans*. *Science* 356, 320–323.

Kocher, S. D. and Paxton, R. J. (2014). Comparative methods offer powerful insights into social evolution in bees. *Apidologie* 45, 289–305.

Kocher, S., Li, C., Yang, W., Tan, H., Yi, S., Yang, X., Hoekstra, H., Zhang, G., Pierce, N. and Yu, D. (2013). The draft genome of a socially polymorphic halictid bee, *Lasioglossum albipes*. *Genome Biol.* 16, 34.

Kucharski, R., Maleszka, J., Foret, S. and Maleszka, R. (2008). Nutritional control of reproductive status in honeybees via DNA methylation. *Science* 319, 1827–1830.

Langer, P., Hogendoorn, K. and Keller, L. (2004). Tug-of-war over reproduction in a social bee. *Nature* 428, 844.

Levis, N. A. and Pfennig, D. W. (2016). Evaluating “plasticity-first” evolution in nature: Key criteria and empirical approaches. *Trends Ecol. Evol.* 31, 563–574.

Levis, N. A., Isdaner, A. J. and Pfennig, D. W. (2018). Morphological novelty emerges from pre-existing phenotypic plasticity. *Nat. Ecol. Evol.* 2, 1289–1297.

Liu, S., Yeh, C. T., Ji, T., Ying, K., Wu, H., Tang, H. M., Fu, Y., Nettleton, D. and Schnable, P. S. (2009). *Mu* transposon insertion sites and meiotic recombination events co-localize with epigenetic marks for open chromatin across the maize genome. *PLoS Genet.* 5, e1000733.

Lyko, F., Foret, S., Kucharski, R., Wolf, S., Falckenhayn, C. and Maleszka, R. (2010). The honey bee epigenomes: Differential methylation of brain DNA in queens and workers. *PLoS Biol.* 8, e1000506.

Michener, C. D. (1974). *The Social Behavior of the Bees: A Comparative Study*. Cambridge, MA: Harvard University Press.

Moczek, A. P., Sultan, S., Foster, S., Ledon-Rettig, C., Dworkin, I., Nijhout, H. F., Abouheif, E. and Pfennig, D. W. (2011). The role of developmental plasticity in evolutionary innovation. *Proc. R. Soc. B Biol. Sci.* 278, 2705.

Morgan, C. (1896). Of modification and variation. *Science* 4, 733–740.

Novick, A. and Weiner, M. (1957). Enzyme induction as an all-or-none phenomenon. *Proc. Natl. Acad. Sci.* 43, 553–566.

Nussey, D. H., Postma, E., Gienapp, P. and Visser, M. E. (2005). Selection on heritable phenotypic plasiticity in a wild bird population. *Science* 310, 304–306.

Osborn, H. F. (1897). Organic selection. *Science* 6, 583–587.

Pereboom, J. J. M., Jordan, W. C., Sumner, S., Hammond, R. L. and Bourke, A. F. G. (2005). Differential gene expression in queen-worker caste determination in bumble-bees. *Proc. R. Soc. B Biol. Sci.* 272, 1145–1152.

Pfennig, D. W., Wund, M. A. and Snell, E. C. (2010). Phenotypic plasticity's impacts on diversification and speciation. *Trends Ecol. Evol.* 25, 459–467.

Pigliucci, M. (2005). Evolution of phenotypic plasticity: Where are we going now? *Trends Ecol. Evol.* 20, 481–486.

Pigliucci, M. (2006). Phenotypic plasticity and evolution by genetic assimilation. *J. Exp. Biol.* 209, 2362–2367.

Price, T. D., Qvarnstrom, A. and Irwin, D. E. (2003). The role of phenotypic plasticity in driving genetic evolution. *Proc. R. Soc. B Biol. Sci.* 270, 1433–1440.

Reeve, H. K. and Gamboa, G. J. (1987). Queen regulation of worker foraging in paper wasps: a social feedback control system (*Polistes fuscatus*, Hymenoptera: Vespidae). *Behaviour* 102, 147–167.

Renn, S. C. P. and Schumer, M. E. (2013). Genetic accommodation and behavioural evolution: Insights from genomic studies. *Anim. Behav.* 85, 1012–1022.

Rissing, S. W. and Pollock, G. B. (1986). Social interaction among pleometrotic queens of *Veromessor pergandei* (Hymenoptera:Formicidae) during colony foundation. *Anim. Behav.* 34, 226–233.

Rittschof, C. C., Bukhari, S. A., Sloofman, L. G., Troy, J. M., Caetano-Anollés, D., Cash-Ahmed, A., Kent, M., Lu, X., Sanogo, Y. O., Weisner, P. A., et al. (2014). Neuromolecular responses to social challenge: Common mechanisms across mouse, stickleback fish, and honey bee. *Proc. Natl. Acad. Sci.* 111, 17929–17934.

Robinson, G. E., Page Jr, R. E., Strambi, A. and Strambi, C. (1989). Hormonal and genetic control of behavioral integration in honey bee colonies. *Science* 246, 109–112.

Robinson, G. E., Page Jr, R. E., Strambi, C. and Strambi, A. (1992). Colony integration in honey bees: mechanisms of behavioral reversion. *Ethology* 90, 336.

Rodgers, A. B., Morgan, C. P., Leu, N. A. and Bale, T. L. (2015). Transgenerational epigenetic programming via sperm microRNA recapitulates effects of paternal stress. *Proc. Natl. Acad. Sci.* 112, 13699–13704.

Romiguier, J., Cameron, S. A., Woodard, S. H., Fischman, B. J., Keller, L. and Praz, C. J. (2016). Phylogenomics controlling for base compositional bias reveals a single origin of eusociality in corbiculate bees. *Mol. Biol. Evol.* 33, 670–678.

Ross, K. and Keller, L. (2002). Experimental conversion of colony social organization by manipulation of worker genotype composition in fire ants (*Solenopsis invicta*). *Behav. Ecol. Sociobiol.* 51, 287–295.

Ruden, D. M. and Lu, X. (2008). *Hsp90* affecting chromatin remodeling might explain transgenerational epigenetic inheritance in *Drosophila*. *Curr. Genomics* 9, 500–508.

Sakagami, S. F. and Maeta, Y. (1984). Multifemale nests and rudimentary castes in the normally solitary bee *Ceratina japonica* (Hymenoptera: Xylocopinae). *J. Kansas Entomol. Soc.* 57, 639–656.

Sakagami, S. and Maeta, Y. (1987). Multifemale nests and rudimentary castes of an "almost" solitary bee *Ceratina flavipes*, with additional observations on multifemale nests of *Ceratina japonica*. *Anim. Soc. Theor. Facts* 55, 391–409.

Sakagami, S. and Maeta, Y. (1989). Compatibility and incompatibility of solitary life with eusociality in two normally solitary bees *Ceratina japonica* and *Ceratina okinawana* (Hymenoptera: Apoidea), with notes on the incipient phase of eusociality. *Japanese J. Entomol.* 57, 417–439.

Santana, S. E. and Dumont, E. R. (2009). Connecting behaviour and performance: The evolution of biting behaviour and bite performance in bats. *J. Evol. Biol.* 22, 2131–2145.

Scharf, M. E., Wu-Scharf, D., Pittendrigh, B. R. and Bennett, G. W. (2003). Caste- and development-associated gene expression in a lower termite. *Genome Biol.* 4, R62.

Schlichting, C. D. and Wund, M. A. (2014). Phenotypic plasticity and epigenetic marking: An assessment of evidence for genetic accommodation. *Evolution* 68, 656–672.

Schulte, C., Theilenberg, E., Muller-Borg, M., Gempe, T. and Beye, M. (2014). Highly efficient integration and expression of piggyBac-derived cassettes in the honeybee (*Apis mellifera*). *Proc. Natl. Acad. Sci.* 111, 9003–9008.

Schwarz, M. P., Richards, M. H. and Danforth, B. N. (2007). Changing paradigms in insect social evolution: insights from halictine and allodapine bees. *Annu. Rev. Entomol.* 52, 127–150.

Shell, W. A. and Rehan, S. M. (2017). Behavioral and genetic mechanisms of social evolution: insights from incipiently and facultatively social bees. *Apidologie* 49, 13–30.

Shen, J. C., Rideout, W. M. and Jones, P. A. (1994). The rate of hydrolytic deamination of 5-methylcytosine in double-stranded DNA. *Nucleic Acids Res.* 22, 972–976.

Sherman, P. W., Lacey, E. A., Reeve, H. K. and Keller, L. (1995). Forum: The eusociality continuum. *Behav. Ecol.* 6, 102–108.

Short, A. K., Yeshurun, S., Powell, R., Perreau, V. M., Fox, A., Kim, J. H., Pang, T. Y. and Hannan, A. J. (2017). Exercise alters mouse sperm small noncoding RNAs and induces a transgenerational modification of male offspring conditioned fear and anxiety. *Transl. Psychiatry* 7, e1114.

Shpigler, H. Y. and Robinson, G. E. (2015). Laboratory assay of brood care for quantitative analyses of individual differences in honey bee (*Apis mellifera*) affiliative behavior. *PLoS One* 10, e0143183.

Shpigler, H. Y., Saul, M. C., Corona, F., Block, L., Ahmed, A. C., Zhao, S. D. and Robinson, G. E. (2017). Deep evolutionary conservation of autism-related genes. *Proc. Natl. Acad. Sci.* 114, 9653.

Siklenka, K., Erkek, S., Godmann, M., Lambrot, R., McGraw, S., Lafleur, C., Cohen, T., Xia, J., Suderman, M., Hallett, M., et al. (2015). Disruption of histone methylation in developing sperm impairs offspring health transgenerationally. *Science* 350, aab2006.

Simola, D. F., Graham, R. J., Brady, C. M., Enzmann, B. L., Desplan, C., Ray, A., Zwiebel, L. J., Bonasio, R., Reinberg, D., Liebig, J., et al. (2015). Epigenetic (re)programming of caste-specific behavior in the ant *Camponotus floridanus*. *Science* 351, aac6633.

Smith, A. R., Wcislo, W. T. and O'Donnell, S. (2003). Assured fitness returns favor sociality in a mass-provisioning sweat bee, *Megalopta genalis* (Hymenoptera: Halictidae). *Behav. Ecol. Sociobiol.* 54, 14–21.

Soucy, S. L. and Danforth, B. N. (2002). Phylogeography of the socially polymorphic sweat bee *Halictus rubicundus* (Hymenoptera: Halictidae). *Evolution* 56, 330–341.

Standage, D. S., Berens, A. J., Glastad, K. M., Severin, A. J., Brendel, V. P. and Toth, A. L. (2016). Genome, transcriptome and methylome sequencing of a primitively eusocial wasp reveal a greatly reduced DNA methylation system in a social insect. *Mol. Ecol.* 25, 1769–1784.

Stark, R. E. (1992). Cooperative nesting in the multivoltine large carpenter bee *Xylocopa sulcatipes* Maa (Apoidea: Anthophoridae): do helpers gain or lose to solitary females? *Ethology* 91, 301–310.

Suzuki, Y. and Nijhout, H. F. (2006). Evolution of a polyphenism by genetic accommodation. *Science* 311, 650–652.

Tenczar, P., Lutz, C. C., Rao, V. D., Goldenfeld, N. and Robinson, G. E. (2014). Automated monitoring reveals extreme interindividual variation and plasticity in honeybee foraging activity levels. *Anim. Behav.* 95, 41–48.

Thompson, G. J. and Oldroyd, B. P. (2004). Evaluating alternative hypotheses for the origin of eusociality in corbiculate bees. *Mol. Phylogenet. Evol.* 33, 452–456.

Toth, A. L., Varala, K., Newman, T. C., Miguez, F. E., Hutchison, S. K., Willoughby, D. A., Simons, J. F., Egholm, M., Hunt, J. H., Hudson, M. E., et al. (2007). Wasp gene expression supports an evolutionary link between maternal behavior and eusociality. *Science* 318, 441–444.

Trible, W., Olivos-Cisneros, L., McKenzie, S. K., Saragosti, J., Chang, N. C., Matthews, B. J., Oxley, P. R. and Kronauer, D. J. C. (2017). *orco* mutagenesis causes loss of antennal lobe glomeruli and impaired social behavior in ants. *Cell* 170, 727–735.

Valtonen, T. M., Kangassalo, K., Pölkki, M. and Rantala, M. J. (2012). Transgenerational effects of parental larval diet on offspring development time, adult body size and pathogen resistance in *Drosophila melanogaster*. *PLoS One* 7, e31611.

Van Buskirk, J. and Relyea, R. a (1998). Selection for phenotypic plasticity in *Rana sylvatica* tadpoles. *Biol. J. Linn. Soc.* 65, 301–328.

Waddington, C. H. (1942). Canalization of development and the inheritance of acquired characters. *Nature* 150, 563–565.

Waddington, C. (1953). Genetic assimilation of an acquired character. *Evolution* 7, 118–126.

Wei, Y., Yang, C. R., Wei, Y. P., Zhao, Z. A., Hou, Y., Schatten, H. and Sun, Q. Y. (2014). Paternally induced transgenerational inheritance of susceptibility to diabetes in mammals. *Proc. Natl. Acad. Sci.* 111, 1873–1878.

West-Eberhard, M. J. (2003). *Developmental plasticity and evolution*. New York: Oxford University Press.

Whitfield, C. W., Cziko, A. M. and Robinson, G. E. (2003). Gene expression profiles in the brain predict behavior in individual honey bees. *Science* 302, 296–299.

Whitfield, C. W., Ben-Shahar, Y., Brillet, C., Leoncini, I., Crauser, D., LeConte, Y., Rodriguez-Zas, S. and Robinson, G. E. (2006). Genomic dissection of behavioral maturation in the honey bee. *Proc. Natl. Acad. Sci.* 103, 16068–16075.

Wilson, E. O. (1971). *The insect societies.* Cambridge: Harvard University Press.

Wilson, E. O. (1985). The sociogenesis of insect colonies. *Science* 228, 1489–95.

Woodard, S. H., Fischman, B. J., Venkat, A., Hudson, M. E., Varala, K., Cameron, S. A., Clark, A. G. and Robinson, G. E. (2011). Genes involved in convergent evolution of eusociality in bees. *Proc. Natl. Acad. Sci.* 108, 7472–7477.

Wray, G., Hoekstra, H., Futuyma, D., Lenski, R., Mackay, T., Schluter, D. and Strassman, J. (2014). Does evolutionary theory need a rethink? No, all is well. *Nature* 514, 161–164.

Yan, H., Opachaloemphan, C., Mancini, G., Yang, H., Gallitto, M., Mlejnek, J., Leibholz, A., Haight, K., Ghaninia, M., Huo, L., et al. (2017). An engineered *orco* mutation produces aberrant social behavior and defective neural development in ants. *Cell* 170, 736–747.

CHAPTER 3

DEVELOPMENTAL TRANSCRIPTOME FOR A FACULTATIVELY EUSOCIAL BEE, *MEGALOPTA GENALIS*[1]

Abstract

Transcriptomes provide excellent foundational resources for mechanistic and evolutionary analyses of complex traits. We present a developmental transcriptome for the facultatively eusocial bee *Megalopta genalis*, which represents a potential transition point in the evolution of eusociality. A *de novo* transcriptome assembly of *Megalopta genalis* was generated using paired-end Illumina sequencing and the Trinity assembler. Males and females of all life stages were aligned to this transcriptome for analysis of gene expression profiles throughout development. Gene Ontology analysis indicates that stage-specific genes are involved in ion transport, cell–cell signaling, and metabolism. A number of distinct biological processes are upregulated in each life stage, and transitions between life stages involve shifts in dominant functional processes, including shifts from transcriptional regulation in embryos to metabolism in larvae, and increased lipid metabolism in adults. We expect that this transcriptome will provide a useful resource for future analyses to better understand the molecular basis of the evolution of eusociality and, more generally, phenotypic plasticity.

[1] The work in this chapter has been adapted from a previously published article: Jones, BM, Wcislo, WT and Robinson, GE. 2015. Developmental transcriptome for a facultatively eusocial bee, *Megalopta genalis*. *G3* 5:2127-2135. Figures and tables have been renumbered. Supplemental files S1, S2 and S3 are available online at http://www.g3journal.org/content/suppl/2015/08/14/g3.115.021261.DC1. Reprinted with permission via the Creative Commons Attribution 4.0 International License: http://www.g3journal.org/content/permissions

Introduction

Transcriptomes provide excellent foundational resources for mechanistic and evolutionary analyses of complex traits in both model and nonmodel organisms. For example, in human disease research, variation in cell- or tissue-specific gene expression has implications for personalized medicine, and transcriptomics has become an attractive approach for cancer diagnosis and therapy choice (Parker et al., 2009; Rosenwald et al., 2002). Evolutionary biologists studying natural populations have increasingly used transcriptomics to bridge the gap between environment and phenotype by revealing context-specific gene expression and the function of novel transcripts and genes (Alvarez et al., 2015). Understanding the extent of gene expression variation can address how responsive a population may be to novel environmental conditions (Oleksiak et al., 2002; Whitehead and Crawford, 2006), and variation in expression can itself be a target of selection (Oleksiak et al., 2002; Whitehead, 2012).

We present a developmental transcriptome for the facultatively eusocial halictid bee *Megalopta genalis*. Our goal is to provide a tool that will enable *M. genalis* to be used in comparative transcriptomic analyses to better understand the evolution of eusociality, one of the most extreme forms of animal developmental phenotypic plasticity. Eusociality evolved independently at least 24 times, nine or more within the hymenopteran insects (Bourke, 2011; Cameron and Mardulyn, 2001; Cardinal et al., 2010; Hines et al., 2007), but absence of extant ancestral lineages prohibits the direct study of eusocial origins.

A promising approach to studying the origins of eusociality is to study incipiently social species across different lineages in a comparative context (Kocher and Paxton, 2014). Particular bee groups are especially well-suited for this task due to the variation in the expression of sociality within and among species. One strikingly diverse group of bees is the family Halictidae, which is a cosmopolitan taxon comprising greater than 4000 species with behavior that ranges from solitary to eusocial (Michener, 1990). Within one subfamily, the Halictinae, at least three independent origins of eusociality have been identified (Danforth, 2002), all of which occurred approximately 20–22 million years ago (Brady et al., 2006). Additionally, there may have been a number of reversions to solitary life among the halictids, suggesting this group is especially flexible and able to transition between solitary and social states (Kocher and Paxton, 2014; Wcislo and Danforth, 1997). Finally, the subfamily Halictinae includes species that are facultatively social, in which females of the same population can produce either solitary or

eusocial nests (Cronin and Hirata, 2003; Eickwort et al., 1996; Packer, 1990, reviewed in Kocher and Paxton, 2014).

One such facultatively eusocial species is *Megalopta genalis*, a Neotropical bee common in the rainforests of the Americas and especially well-studied on Barro Colorado Island in Panama (Wcislo et al., 2004). Foundress females of this species can produce either solitary nests, with only male offspring in the first brood, or small eusocial nests with at least one daughter worker in the first generation (Kapheim et al., 2013; Smith et al., 2003; Wcislo et al., 2004). Both solitary and eusocial nests may produce a mix of dispersing males and females in later generations, and differences in sex ratio are not due to mating status of females because all reproductive females are mated (Kapheim et al., 2012). Instead, flexibility in nest sociality is a result of larval and adult environmental influences (Kapheim et al., 2013; Smith et al., 2003), and may represent a transition point in the evolution of eusocial insects. If the phenotypic flexibility present in *M. genalis* captures an evolutionary transition in social behavior, then understanding the mechanisms of this flexibility may open a window into the origins of eusociality. However, *M. genalis* is currently limited as a model for social transitions due to the lack of resources for studies of gene expression or genetic underpinnings of social flexibility.

As the first step toward using transcriptomic analyses of *M. genalis* to better understand the role of developmental phenotypic plasticity in the origins of eusociality, we sequenced and assembled a developmental transcriptome. We used this transcriptome to conduct a preliminary survey of the extent of plasticity in gene expression in *M. genalis* across development and identified molecular pathways with highly plastic gene expression. In the future, we expect this reference transcriptome to be useful in studies of gene expression in *M. genalis* as well as in a comparative framework with other social insects in studying the developmental origins of eusociality.

Materials and Methods

Sample collection and tissue preparation

Collections were made on Barro Colorado Island (BCI) in the Republic of Panama, a 1500-hectare island in Lake Gatun formed during construction of the Panama Canal. *Megalopta* species are nocturnal insects active during the dry season, with densities of approximately $5x10^{-3}$ nests per square meter (Wcislo et al., 2004; Wolda and Roubik, 1986). Some individuals were

collected from natural nests and then placed into liquid nitrogen. Other individuals were collected as larvae or pupae from natural nests and then reared through the adult stage at ambient temperature in an outdoor enclosure prior to liquid nitrogen freezing. A subset of females was placed into observation nests (as described in Kapheim et al., 2011) after eclosion and collected after nest construction and egg laying had begun (rearing and age information where known is provided in File S3). Differential effects of rearing condition were not apparent from clustering of samples based on gene expression variation (Figure 3.1B). A total of four eggs, eight larvae, eight pupae, and 20 adults were used for sequencing. Pupal and adult samples were balanced for sex, and the sexes of earlier life stages were determined postsequencing as explained below (see *Sex determination of preadult stages*).

For adult samples, whole brains were dissected from the head capsule on wet ice following 16-hr incubation at 22º in RNA-later ICE (Life Technologies), with frons and cuticle around postocciput removed prior to incubation. Abdomens were similarly incubated in RNA-later ICE prior to removal of gut tissue and Dufour's glands. Preadult life stages were not dissected. Many tissue types were included for sequencing to maximize the number of transcripts captured for transcriptome assembly; however, a caveat of the data is that tissue- or cell-type resolution is lost. Total RNA was extracted using QIAGEN RNeasy Mini Kits [treated with DNase (QIAGEN) to remove genomic DNA] and quality was confirmed with Bioanalyzer RNA Pico chips (Agilent) prior to library preparation.

Library preparation for RNA-sequencing (RNA-seq)

Poly-A RNA was enriched from 0.6 to 1.0 mg total RNA using NEXTflex Poly(A) Beads from Bioo Scientific. Strand-specific cDNA libraries were prepared using the Bioo Scientific NEXTflex Directional RNA-Seq Kit (dUTP Based) for Illumina following manufacturer instructions with a 12-min fragmentation time and 15 PCR cycles. Libraries were barcoded with Bioo Scientific adaptors so they could be pooled for sequencing. For the 20 adult samples, two libraries were created per individual: one for whole brain and one for abdominal tissue. In total, 60 libraries were created for the 40 individuals in the study. Library concentrations were quantified using a Qubit dsDNA High Sensitivity Assay Kit (Life Technologies), and library size was assessed using a Bioanalyzer High Sensitivity DNA chip (Agilent). Libraries were pooled into four groups at equal concentration and diluted to final pooled concentrations of 10 nM. Library pools were quantified using the Illumina-compatible kit and KAPA standards for real-

time PCR by the W. M. Keck Center for Comparative and Functional Genomics at the Roy J. Carver Biotechnology Center (University of Illinois).

Sequencing and preassembly read processing

Paired-end sequencing was performed on an Illumina HiSeq2000 at the W. M. Keck Center. Fifteen libraries were sequenced on each of four lanes, resulting in over 1.6 billion reads (averaging 27.12 million per library). Quality was assessed using FASTQC (v. 0.10.1), and read trimming was performed with Trimmomatic (v. 0.32) to remove low-quality reads and remaining adapter sequences. Ninety-eight percent of reads passed quality and adapter trimming across all samples. For assembly, reads were concatenated into single files prior to running a digital normalization to a maximum coverage of 50x. This normalization reduced the number of input reads by nearly 90%, dramatically reducing computer processing time, and was expected to result in little to no loss of transcript information (Brown et al., 2012).

Trinity assembly

An initial *de novo* assembly was performed using Trinity (v20140413) and included all 60 libraries. This assembly resulted in nearly 197,000 genes and 256,000 transcripts, a highly unrealistic number of transcripts given what we know from other bee genomes (an average of 13,616 genes for 10 bee genomes sequenced) (Kapheim et al., 2015). Mapping these transcripts to the closest related genome available, that of *Lasioglossum albipes* (Kocher et al., 2013), revealed that many transcripts and genes were mapping to the same loci. To reduce the complexity of reads for assembly, we took advantage of a set of five related individuals within our sample group (Appendix Figure A.1). RNA from these individuals was used to make nine libraries, which were sequenced and resulted in a combined total of 142 million reads. As before, digital normalization was applied to 50x using Jellyfish. Trinity (v20140413) was again used for assembly and, as expected, the number of genes and transcripts was reduced to 75,206 genes and 102,303 transcripts. Additionally, the contig N50 and other measures of assembly continuity and quality improved (contig N50: increased from 1056 to 2057 bp; mean contig length: increased from 658.19 to 868.98 bp). Finally, the percentage of reads mapping (methods described in *Read alignment and abundance estimation*) to this assembly was higher than the previous assembly (84% vs. 82% average per sample). This assembly was therefore used for the remaining analyses presented. Assembly statistics are reported in Table 3.1, where gene-level metrics are based on the longest isoform per gene.

Assembled transcripts were screened against the NCBI nonredundant (nr) database using BLASTX with an e-value threshold of 1e-5. Of the 102,303 transcripts, 37.52% had a significant hit to the nr database; 96.61% of these hits were to insects, and 95.34% of hits were to hymenopteran species. Although nr is one of the most complete databases of sequence information available to the public, it does not yet contain information from five bee genomes that have been recently sequenced (Kapheim et al., 2015). Using a BLASTN against a custom database consisting of all 10 bee genomes that have been sequenced to date (Appendix Figure A.2), 55.20% of assembled transcripts (and 45.71% of genes) mapped to at least one location in at least one of the 10 bee genomes with an e-value threshold of 1e-5.

Completeness of the assembly was assessed using two sets of information from the 10 bee genomes. The first comparison set included 5855 single copy orthologs across all 10 genomes (Kapheim et al., 2015). All assembled transcripts were used as queries in a BLASTN (maximum e-value of 10e-3) against this set of orthologous genes. Similarly, all transcripts were mapped against each of the 10 bee genomes, and the percentage of unique genes with a transcriptome hit is shown in Appendix Figure A.2 (range: 37.17–84.56%).

Read alignment and abundance estimation

Quality- and adapter-trimmed reads for all 60 libraries were aligned to the transcriptome using the align_and_estimate_abundance.pl script in the Trinity r20140413 toolkit, which uses Bowtie (version 0.12.7) for alignment and RSEM (version 1.2.10) for estimating transcript abundance. An average of 84% of reads per sample (min: 74.14%; max: 87.95%) reported at least one alignment to the transcriptome. For all downstream analyses, only read counts at the putative gene level (not transcript level) were used.

Sex determination of preadult stages

The Haplotype Caller within the Genome Analysis Toolkit (GATK) was used to predict sex for the preadult stages. Because males are haploid, the number of confidently called SNPs for a male should be small compared to the number of SNPs found in diploid females. As a proof of concept, 12 adult individuals (of known sex) were run through Haplotype Caller. Female individuals (n = 6) had 23,857 ± 3428 SNPs called (suggestive of heterozygous loci) while males (n = 6) had only 3509 ± 1564 SNPs using the same filtering criteria (these SNPs could be the result of paralogous gene sequences, sequencing errors, or assembly artifacts). The sexes of all eggs and larvae were assigned based on the number of SNPs called using Haplotype Caller, with

all predicted females having greater than 12,500 SNPs and all males having less than 3800 called SNPs. Three of the four eggs and six of the eight larvae were predicted to be female based on these criteria.

Principal component analysis

Broad clustering of gene expression profiles for all individuals was conducted using a Principal Component Analysis (PCA) of TMM-normalized (Robinson and Oshlack, 2010) FPKM values obtained using RSEM within the Trinity script abundance_estimates_to_matrix.pl. The PtR.pl script packaged with Trinity, which utilizes a number of plotting functions in R, was used to produce the plot and heatmap presented in Figure 3.1. Genes with less than 10 FPKM counts across the 60 libraries were excluded prior to clustering, and data were log2 transformed prior to PCA. The most variable 100 genes (based on extreme eigenvalues) for each of the first four principal components (400 genes total) are shown in the heatmap and clustering dendrogram. This analysis provides a visualization of gene expression based on the most variably expressed genes across all samples.

Developmental dynamics of gene expression

We conducted a preliminary survey of gene expression changes throughout development using the R package maSigPro, which uses a GLM regression approach to find clusters of genes significantly differentiated through time (i.e., across life stages; a linear step-up Benjamini-Hochberg false discovery rate procedure was used, with corrected $P < 0.05$ for all genes) (Conesa et al., 2006). The 75,206 *M. genalis* assembled genes were filtered to include only those with least 1 count per million (CPM) in at least two samples, resulting in 22,315 genes for analysis. To avoid inappropriate grouping of potentially distinct groups of genes, we initially used the maximum number ($n = 9$) of clusters for maSigPro, followed by paring down to six clusters based on similar expression patterns of three pairs of clusters (the original nine clusters are shown in Appendix Figure A.3). A design matrix was formed that described the life stage and tissue (egg, larva, pupa, adult-abdomen, adult-brain) of each individual. For each cluster of genes, the median expression value of those genes for each individual is calculated, and this median for each individual is then averaged across samples for visualization in Figure 3.2. The analysis was repeated excluding males, and results looked very similar (Appendix Figure A.4).

For each life stage, we identified genes more highly expressed in that life stage (for adults considering abdomens and brains separately) than in any other life stage using differential gene

expression analysis with edgeR (Robinson et al., 2010). The filtered set of 22,315 genes with CPM ≥1 in at least two samples was used for analysis. Count data (obtained from RSEM as discussed above) were normalized by library size and library composition (TMM) in edgeR. Dispersion was estimated across samples using the estimateGLMTagwiseDisp function in R (following estimateGLMTrendedDisp) such that estimates were squeezed toward the trended dispersion values with a prior degrees of freedom value of 20 (McCarthy et al., 2012). Raw *P* values from each test were commonly corrected using the p.adjust function in R using the Benjamini-Hochberg ("FDR") method. Genes that were more highly expressed in one life stage compared to all other stages (using an FDR-corrected *P* of 0.05 as the significance cutoff) were functionally annotated using PANTHER, and statistical overrepresentation tests were conducted on those lists relative to the reference set of 22,315 genes used in the edgeR analysis.

To assess changes associated with transitions between life stages, we used PANTHER overrepresentation tests on differential expression lists obtained from edgeR such that each life stage was compared with life stages directly preceding or following that stage (the pupal stage was compared with the adult brain and adult abdomen samples separately). For example, genes that were more highly expressed in larvae compared with eggs were compared to the reference 22,315 genes to test for overrepresentation of GO-Slim and PANTHER protein categories as described below.

Functional annotation of genes with PANTHER

TransDecoder (r20131110, packaged with Trinity r20140413) was used to identify candidate coding regions within assembled *M. genalis* transcripts. The predicted peptides for the 22,315 genes that passed the minimum expression threshold were used as input to PANTHER (pantherScore1.03, library version 9.0) to identify protein family domains. A statistical overrepresentation test with Bonferroni correction for multiple testing was used to identify biological processes and protein classes that were overrepresented in the overexpressed gene lists for each life stage relative to the reference set of 22,315 genes. These analyses were conducted using the Gene List Analysis tools available on the PANTHER website (pantherdb.org) (Mi et al., 2013).

Data availability

File S1 contains enriched terms from PANTHER analysis of *M. genalis* gene clusters. File S2 lists enriched terms from PANTHER analysis of genes overexpressed in each life stage.

File S3 provides rearing and age information for sequenced individuals. Sequencing reads used for transcriptome assembly and differential expression analysis have been deposited in the short read archive (SRA, NCBI) under the accession number SRP057750.

Results and Discussion

Reference transcriptome assembly statistics

To reduce the complexity of reads for assembly, we took advantage of a set of five related individuals within our sample group (Appendix Figure A.1). RNA from these individuals was used to make nine libraries (RNA from brains and abdomens in separate libraries for adults), which were sequenced and resulted in a combined total of 142 million reads; 75,206 Trinity components (hereafter referred to as genes) were assembled from these reads, yielding a contig N50 of 2057 bp. Additional assembly statistics are presented in Table 3.1. Assembly completeness was assessed in two ways using data from genome sequences of 10 bee species (Kapheim et al., 2015): (1) we determined the presence or absence of 5855 single copy orthologs identified as common to all 10 bees and (2) we compared our list of putative genes in *M. genalis* with the genes identified in each of the 10 sequenced bee genomes. Ninety-seven percent of genes within the orthologous gene set were found in the *M. genalis* transcriptome, indicating that sequencing depth and assembly parameters were sufficient to capture nearly all highly conserved transcripts. The two species with the highest percentage of genes with sequence homology to transcripts in the transcriptome are *Lasioglossum albipes* (75.33%), the species most closely related to *Megalopta genalis*, and *Apis mellifera* (84.56%), the species with the most thoroughly annotated genome (Elsik et al., 2014; Honeybee Genome Sequencing Consortium, 2006). The phylogeny of the 10 bee species used for comparison and the percentage of genes from each of the species with a homologous *M. genalis* contig are shown in Appendix Figure A.2.

Read alignment and abundance estimation

RNA-seq libraries sequenced from 40 individuals spanning all life stages and both sexes were aligned to the reference assembly using Bowtie, and abundance estimation was conducted using RSEM. An average of 84% of quality-trimmed reads mapped to the assembly from each library (range: 74.14–87.95%); 22,315 genes (approximately 30% of total genes) had an expression value of at least 1 CPM in at least two samples. This stringent reads-mapping cutoff was used to give high confidence that the genes in this set are not sequencing artifacts or a result

of assembly errors, and the 22,315 genes above the 1 CPM cutoff were used for all differential expression and functional analyses. This number is similar to that found in a de novo assembly of the paper wasp *Polistes canadensis* (26,284 isogroups) (Ferreira et al., 2013), but much lower than the number of transcripts (358,709) analyzed in the *de novo* assembly of the small carpenter bee, *Ceratina calcarata* (Rehan et al., 2014).

Survey of gene expression

Principal component analysis: PCA was conducted to identify broad patterns of gene expression and genes that best discriminate sex and life stage groups. The first four principal components (PCs) explained a total of 35.7% of the variance in expression (16.7%, 8.8%, 6.3%, and 3.9%, respectively). Clustering of samples based on the first two PCs is shown in Figure 3.1A, and a heatmap of 400 genes with highly variable expression (100 most extreme genes from each of the first four principal components) is shown in Figure 3.1B.

As evident from the heatmap in Figure 3.1B, the most variable genes separated adult tissues from preadult life stages. Within adults, males and females clustered separately, with one exception (highlighted with arrows in Figure 3.1B) being a single female clustering with males in terms of both brain and abdominal gene expression patterns. This female is unique in that she was newly eclosed when collected, while all other females were collected from nests and are mature adults (potentially months old; File S1). Males only remain in the natal nest for a few days prior to dispersal, thus all males in this study are very likely less than a few days old. Clustering of a very young female with the males suggests that young adults of *M. genalis* may be similar in gene expression regardless of sex, which is in contrast with the obligately eusocial honey bee, which has drastic differences in brain gene expression between sexes even at 1 d old (Zayed et al., 2012). Intriguingly, males and females of *M. genalis* are much more similar in terms of their size and gross morphology than are females and drones of *A. mellifera*, and perhaps the gene expression patterns in the two species reflect the respective levels of morphological differentiation found between sexes. Because the focus of this study was to collect diverse samples for the reference transcriptome, constraints on collecting age-matched males for comparison with adult females prevent a powered analysis of sex differences in gene expression in the adult stage.

For preadult stages, sex was much less predictive for sample clustering and all pupae were completely intermixed with respect to sex. Egg and larval stages also showed little

differentiation based on sex, although sample sizes for males of these stages were small (only one male egg and two male larvae). One female egg clustered more closely with larvae than the other eggs. Because the time of egg laying is unknown for these individuals, one possible explanation is that this egg was close to hatching into the first larval instar. The lack of differentiation between sexes in these early life stages may also reflect differences in the specific developmental time points of the individuals. It is interesting to note that despite the variation included within each life stage (e.g., multiple embryonic stages, different larval instars) and the very small sample sizes, life stages were still strongly differentiated in terms of gene expression.

Developmental dynamics of gene expression: We utilized three separate analyses: (1) a GLM regression approach to find clusters of genes differentiated through time (maSigPro) (Conesa et al., 2006); (2) edgeR to find genes most highly expressed in each life stage relative to all others; and (3) edgeR to find genes differentially expressed in each transition between life stages (Robinson et al., 2010).

Genes were annotated for biological functions using the PANTHER database, and statistical overrepresentation tests were conducted on each gene list relative to the reference set of 22,315 genes with CPM ≥1 in at least two samples (Mi et al., 2013). To describe major differences among genes in each list, we identified the top unique overrepresented biological processes and PANTHER protein classes (Bonferroni-corrected P value, all $P \leq 0.05$) for each list. All significantly overrepresented and underrepresented GO-Slim Biological Processes and PANTHER Protein Classes for each analysis are listed in File S1 and File S2. Results from the three analyses were similar, adding confidence to the signal and providing a broad view of gene expression during each developmental stage of *M. genalis*.

The embryonic stage was dominated by signatures of transcription and DNA binding (Table 3.3), similar to genes in Cluster 6 (Figure 3.2, Table 3.2). This likely reflects the extensive pattern formation and regional specification that occurs during the embryonic stage, with transcription regulation factors such as Wnt establishing body axis patterning and cell fate during the embryonic stage (Cadigan and Nusse, 1997). Cluster 6 was also enriched for a number of biological processes and protein classes related to RNA processing and RNA metabolism (Table 3.2). These genes may be responsible for the rapid differentiation of cell types during the embryonic stage of insects (Shields et al., 1975).

In transitioning to the larval stage, an increase in many metabolic functions and a

decrease in transcriptional regulation were observed (Figure 3.3). High expression of genes enriched for metabolic processes may contribute to the rapid growth through the larval instars as individuals consume pollen provisions in their cells. This result is similar to what has previously been found comparing gene expression between larval and adult ants of *Camponotus festinatus*, with protein metabolism genes highly expressed during the larval stage (Goodisman et al., 2005). Elevated expression of metabolic and storage protein genes has also been reported in larvae of the bumble bee, *Bombus terrestris* (Colgan et al., 2011).

Larvae also showed the highest expression of genes involved in many enzymatic functions, including dehydrogenase and hydrolase protein classes (Table 3.3). This again highlights the turnover of metabolites during the larval growth stages of this insect. Finally, the ubiquitin proteasome pathway is more active in the larval stage compared with the pupal stage, reminiscent of the caste-specific expression of ubiquitin-related genes in larvae of the honey bee (Barchuk et al., 2007; Chen et al., 2012; Humann and Hartfelder, 2011).

Genes more highly expressed in the pupal stage relative to other stages were largely unclassified based on conserved protein domains (Table 3.3). Because the PANTHER protein database currently includes only two insects (*Drosophila melanogaster* and *Anopheles gambiae*), particular protein families important for insect metamorphosis may be underrepresented in PANTHER. Further, bee-specific protein families are absent from the database. However, structural proteins were enriched in the pupal stage, highlighting the extensive physical rearrangements of tissues occurring during metamorphosis. Relative to the larval stage, pupae also showed an increase in neurological system process genes and genes involved in cell–cell signaling, perhaps related to the reorganization of nervous tissues during the pupal stage (Technau and Heisenberg, 1982; White and Kankel, 1978).

Cluster 5 contained 131 genes expressed throughout egg, larval, and pupal stages but lowly expressed in adults. This cluster was enriched for genes involved in cellular and developmental processes, as well as cell adhesion (Table 3.2). This pattern of expression for adhesion genes has also been documented in *Drosophila melanogaster*, with relatively little expression in adults, but expression throughout earlier life stages (Arbeitman et al., 2002). Cluster 5 was also enriched for the cytoskeletal regulation by Rho GTPase pathway (P = 0.00112). The Rho family of GTPases is known to regulate a number of cellular functions important for cell shape, motility, and adhesion, as well as progression through the cell cycle

(Narumiya and Morii, 1993; Ridley, 1995). The expression pattern of these genes as seen in Figure 3.2 suggests that factors influencing growth and cellular organization, while necessary during developmental phases, are of diminished importance in postmitotic adult tissue.

In contrast to genes within Cluster 5, a number of genes had expression restricted to the adult stage (e.g., Clusters 1 and 3). Adult abdominal tissues had significantly higher expression of many genes related to lipid and fatty acid metabolism compared with other life stages (Table 3.3). Insect abdominal fat bodies play a critical role in the storage and utilization of energy (reviewed in Arrese and Soulages, 2010), and thus it is reasonable that lipid metabolism dominates the signal coming from abdominal overexpressed genes. Genes within Cluster 3, which are expressed in both brain and abdominal tissue of adults, are enriched for the DNA photolyase protein class. DNA photolyases are known to repair DNA damage caused by UV radiation (Sancar, 1994), and may be playing a role in mitigating the effects of light exposure in *M. genalis* adults with their exceptionally sensitive eyes (Greiner et al., 2004).

The adult brain showed a strong signal of temporally- and spatially-restricted gene expression, with nearly 5000 genes expressed more highly in the adult brain than in any other life stage or tissue (Table 3.3). In both the ant *Camponotus festinatus* and in *Drosophila melanogaster*, genes highly expressed in the adult stage (including the adult brain) show a greater diversity of functional categories relative to genes more highly expressed in earlier stages (Goodisman et al., 2005).

The cluster of genes showing brain-restricted expression (Cluster 1, Figure 3.2) were also enriched for the muscarinic acetylcholine 1 and 3 signaling pathway ($P = 0.000441$). Muscarinic acetylcholine signaling has been implicated in nestmate recognition, an important feature of social behavior and potential prerequisite for social evolution (Ismail et al., 2008). In addition, this signaling pathway is important for foraging-dependent changes in the structure of the mushroom bodies in honey bees (Ismail et al., 2006). As shown in Figure 3.2, genes in Cluster 1 were more highly expressed in the adult brain of male individuals. Because males in this study were quite young, the inferred greater muscarinic acetylcholine receptor signaling in these individuals suggests that neuron outgrowth may be particularly enriched in early adult life of *M. genalis*, similar to what has been shown for honey bees (Fahrbach et al., 1998) and bumble bees (Jones et al., 2013). However, acetylcholine is one of the most common excitatory neurotransmitters in the insect brain (Gerschenfeld, 1973; Pitman, 1971), and thus could be

involved in numerous other functions in adults.

Conclusions

The ability to develop genomic resources for nonmodel organisms greatly improves our ability to use naturally occurring variation to answer important questions in evolutionary biology (Domingues et al., 2012; Martin et al., 2012). In this study, we presented a comprehensive transcriptome of development in a facultatively eusocial bee, *Megalopta genalis*, an important emerging model for understanding potential precursors to obligate eusociality among social insects. In this early broad survey of gene expression, we found a number of gene clusters with dynamic and/or temporally specific expression profiles throughout development in this bee. Many of these clusters are functionally enriched for particular classes of protein families, and thus open the door to more in-depth gene expression analyses and examinations of how the biological processes implicated here contribute to the phenotypic plasticity exhibited by *M. genalis*. Transitions between life stages of *M. genalis* display striking changes in the functional categories of expressed genes, and life stages show distinct signatures of molecular functions. These results provide a foundation for future studies of transcriptomics in *M. genalis*, as well as more in-depth analyses of gene expression plasticity in facultatively social systems.

Acknowledgments

I would like to thank William Wcislo for his contributions as co-author. Collections were aided by Betzi Perez and Ernesto Gomez-Perez. I thank members of the Robinson lab for helpful discussion and comments on this work, and Troy Comi for help with scripts for data analysis. Matthew Hudson provided helpful feedback during transcriptome assembly and annotation. This work was funded through the University of Illinois NSF IGERT grant (NSF DGE IGERT-1069157, PI Suarez UIUC) and fellowship support from the Smithsonian Tropical Research Institute (STRI). Additional support was provided by STRI research funds (to William Wcislo) and the NIH Director's Pioneer Award (DP1 OD006414 to Gene Robinson). I thank the National Authority for the Environment of the Government of Panama for export permits (no. SEX/A-53-13). Much of the computing work for this research was conducted on the National Science Foundation funded MRI-R2 project #DBI-0959894 (Data Intensive Academic Grid).

Figures and Tables

Figure 3.1

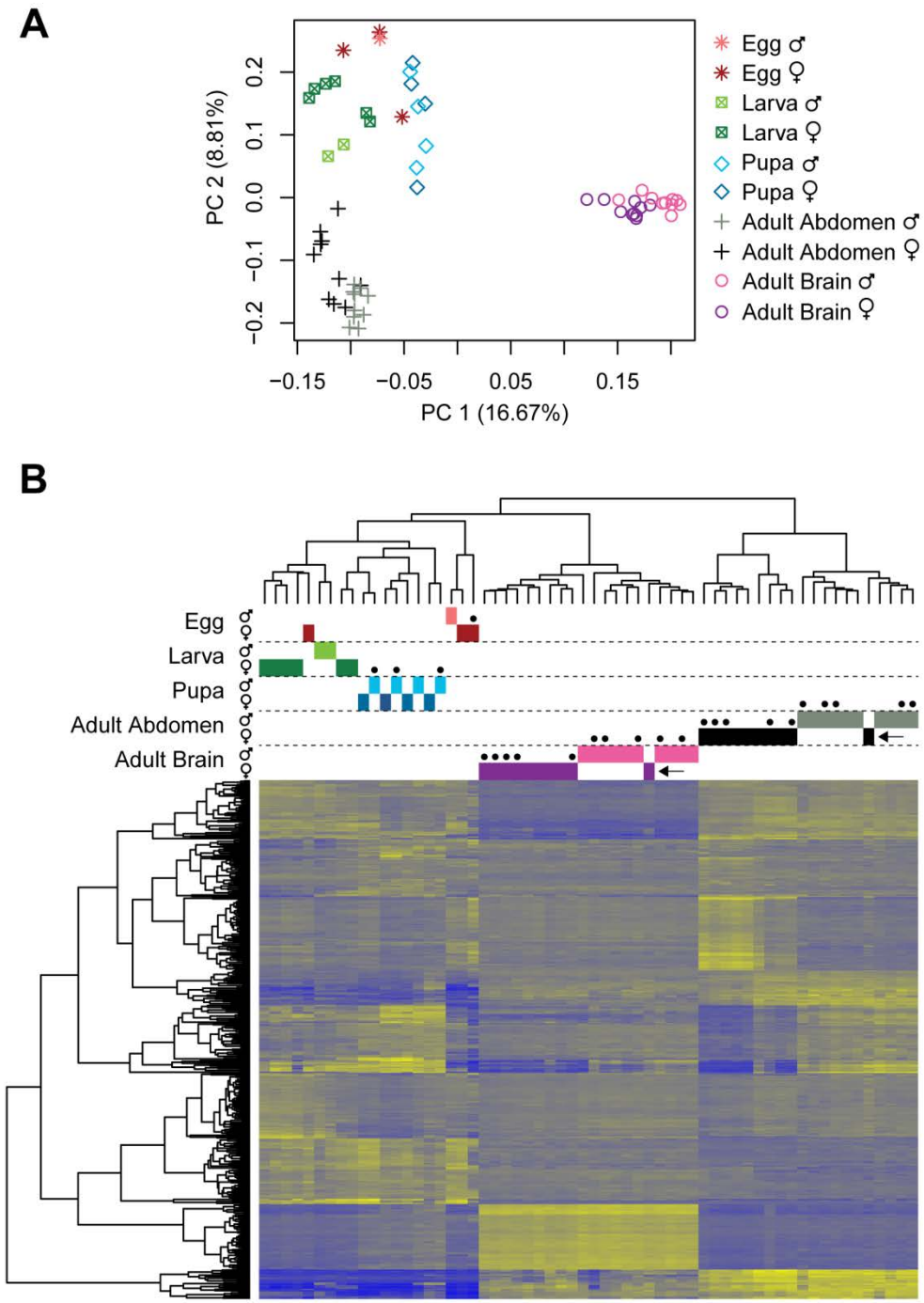


Figure 3.1. Clustering of individuals based on overall patterns of gene expression.

(Figure 3.1 caption continued)

(A) PC plot for first two principal components and (B) heatmap and clustering of 400 genes with the most variable gene expression across samples (100 most variable genes from each of the first four principal components). Each row of the heatmap represents a single gene, and genes are clustered based on expression similarity. Dots above heatmap indicate individuals reared in the laboratory or placed in observation nests prior to collection (as opposed to collected from natural nests in the field). Blue and yellow colors in the heatmap correspond to low or high relative gene expression, respectively. Arrows refer to individual discussed in Principal component analysis section of Results and Discussion.

Figure 3.2

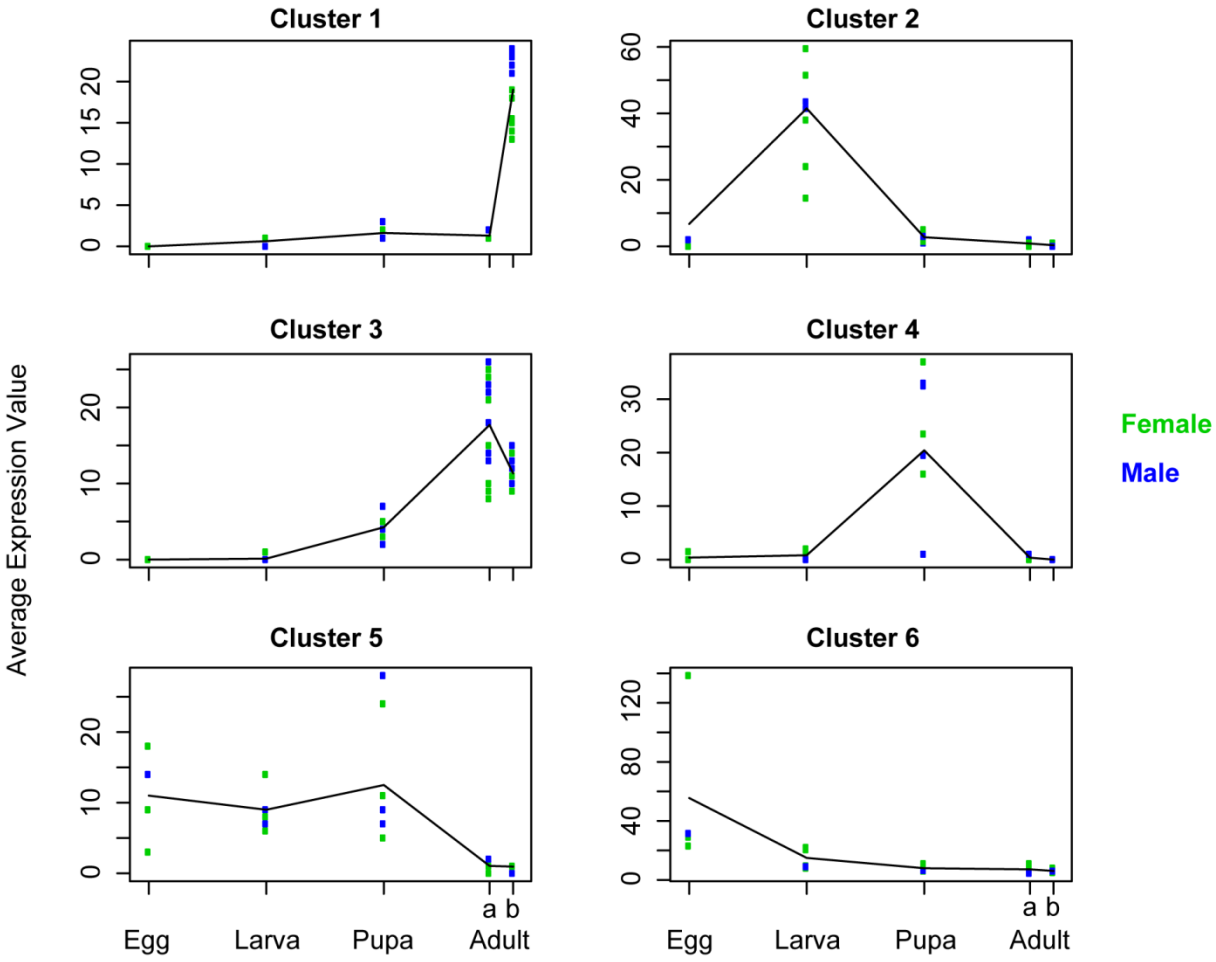


Figure 3.2. Clusters of genes with similar gene expression patterns throughout development. Each symbol represents the median expression for all genes within the cluster for one individual, and lines connect the average expression value across individuals for each life stage. Along the x-axis, “a” and “b” refer to adult abdominal and brain tissues, respectively.

Figure 3.3

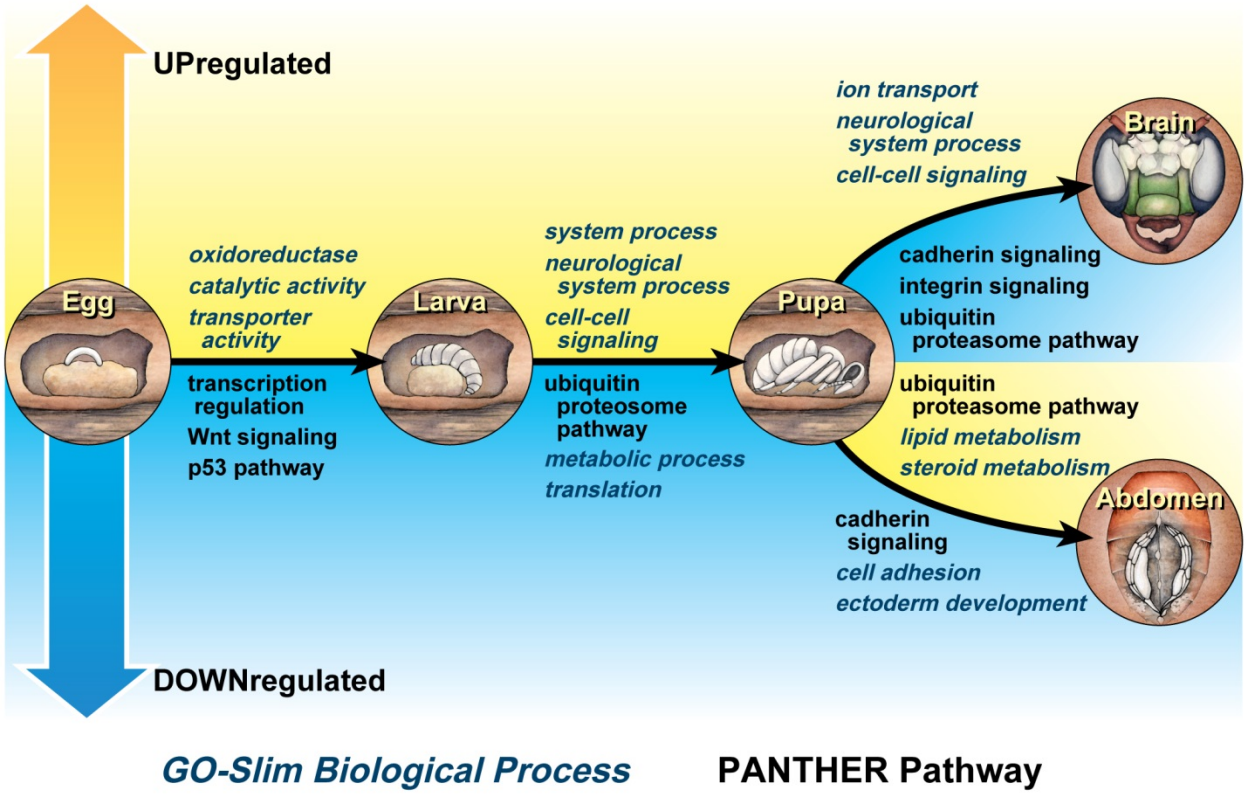


Figure 3.3. Functional annotation of genes showing differences in expression associated with pairwise transitions between life stages.

Terms above transition arrows indicate genes that are more highly expressed in the life stage to the right of the arrow, while terms below the arrow indicate genes that are more highly expressed in the life stage to the left of the transition arrow. Terms in black are PANTHER Pathways, while blue italicized terms are GO-Slim Biological Processes. All terms listed are statistically overrepresented with a Bonferroni-corrected $p < 0.05$. Artistic renderings of different life stages and tissue types are not representative of every sample included in the analysis, and only represent one particular life stage, sex, or tissue. Drawings by Julie Himes.

Table 3.1. Summary of assembly statistics.

Category	Number							
	200–499 bp	*500–999 bp*	*1–1999 bp*	*≥2 kbp*	**Total Number**	**Mean Length (bp)**	**N50 (bp)**	**Total Nucleotides**
Transcripts	51,469	15,504	12,567	22,763	102,303	1390.39	3351	142,241,080
Genes	48,270	11,958	6728	8250	75,206	868.98	2057	65,352,587

Table 3.2. The top four unique overrepresented terms (all Bonferroni-corrected $p < 0.05$) are shown corresponding to gene clusters shown in Figure 3.2.

Cluster	# of Genes	GO-Slim Biological Process	PANTHER Protein Class
1	870	Ion transport, localization, cell-cell signaling, steroid metabolic process	Transporter, oxygenase
2	260	Proteolysis, metabolic process, protein folding, lipid metabolic process	Serine protease, chaperonin, hydrolase, storage protein
3	63	None	DNA photolyase
4	160	Unclassified	Unclassified
5	131	Cellular process, developmental process, cellular component movement, cell adhesion	Receptor, cell adhesion molecule, cadherin, tubulin
6	114	Nucleobase-containing compound metabolic process, mRNA processing, RNA metabolic process, RNA splicing	Nucleic acid binding, RNA binding protein, mRNA processing factor, mRNA splicing factor

Table 3.3. The number of genes significantly overexpressed in each life stage produced from pairwise comparisons of gene expression, as well as the three GO-Slim biological processes and PANTHER protein classes most highly overrepresented for each gene list (all Bonferroni-corrected $p < 0.05$).

Life Stage	# of Genes	GO-Slim Biological Process	PANTHER Protein Class
Egg	837	Nucleobase-containing compound metabolic process, DNA-dependent transcription, transcription from RNA polymerase II promoter	DNA binding protein, transcription factor, nucleic acid binding
Larva	717	Metabolic process, primary metabolic process, lipid metabolic process	Oxidoreductase, dehydrogenase, hydrolase
Pupa	793	Unclassified	Structural protein, unclassified
Adult abdomem	2392	Lipid metabolic process, fatty acid metabolic process, steroid metabolic process	Oxidoreductase, oxygenase, acyltransferase
Adult brain	4924	Neurological system process, system process, cell-cell signaling	Ion channel, ligand-gated ion channel, acetylcholine receptor

References

Alvarez, M., Schrey, A. W. and Richards, C. L. (2015). Ten years of transcriptomics in wild populations: What have we learned about their ecology and evolution? *Mol. Ecol.* 24, 710-725.

Arbeitman, M. N., Furlong, E. E. M., Imam, F., Johnson, E., Null, B. H., Baker, B. S., Krasnow, M. A., Scott, M. P., Davis, R. W. and White, K. P. (2002). Gene expression during the life cycle of *Drosophila melanogaster*. *Science* 297, 2270–2275.

Arrese, E. L. and Soulages, J. L. (2010). Insect fat body: Energy, metabolism, and regulation. *Annu. Rev. Entomol.* 55, 207–225.

Barchuk, A. R., Cristino, A. S., Kucharski, R., Costa, L. F., Simões, Z. L. P. and Maleszka, R. (2007). Molecular determinants of caste differentiation in the highly eusocial honeybee *Apis mellifera*. *BMC Dev. Biol.* 7, 70.

Bourke, A. F. G. (2011). *Principles of Social Evolution*. Oxford: Oxford University Press.

Brady, S. G., Sipes, S., Pearson, A. and Danforth, B. N. (2006). Recent and simultaneous origins of eusociality in halictid bees. *Proc. R. Soc. B Biol. Sci.* 273, 1643–1649.

Brown, C. T., Howe, A., Zhang, Q. and Pyrkosz, A. B. (2012). A reference-free algorithm for computational normalization of shotgun sequencing data. *arXiv* 1203, 4802.

Cadigan, K. M. and Nusse, R. (1997). *Wnt* signaling: A common theme in animal development. *Genes Dev.* 11, 3286–3305.

Cameron, S. A. and Mardulyn, P. (2001). Multiple molecular data sets suggest independant origins of highly eusocial behavior in bees (Hymenoptera: Apinae). *Syst. Biol.* 50, 194–214.

Cardinal, S. C., Paxton, R. J., Field, J., Bridge, C. and Soro, A. (2010). Genetic differentiation across the social transition in a socially polymorphic sweat bee, *Halictus rubicundus*. *Mol. Ecol.* 19, 3351–3363.

Chen, X., Hu, Y., Zheng, H., Cao, L., Niu, D., Yu, D., Sun, Y., Hu, S. and Hu, F. (2012). Transcriptome comparison between honey bee queen- and worker-destined larvae. *Insect Biochem. Mol. Biol.* 42, 665.

Colgan, T. J., Carolan, J. C., Bridgett, S. J., Sumner, S., Blaxter, M. L. and Brown, M. J. F. (2011). Polyphenism in social insects: Insights from a transcriptome-wide analysis of gene expression in the life stages of the key pollinator, *Bombus terrestris*. *BMC Genomics* 12, 623.

Conesa, A., Nueda, M. J., Ferrer, A. and Talón, M. (2006). maSigPro: A method to identify significantly differential expression profiles in time-course microarray experiments. *Bioinformatics* 22, 1096–1102.

Cronin, A. L. and Hirata, M. (2003). Social polymorphism in the sweat bee *Lasioglossum (Evylaeus) baleicum* (Cockerell) (Hymenoptera, Halictidae) in Hokkaido, northern Japan. *Insectes Soc.* 50, 379–386.

Danforth, B. N. (2002). Evolution of sociality in a primitively eusocial lineage of bees. *Proc. Natl. Acad. Sci.* 99, 286–290.

Domingues, V. S., Poh, Y. P., Peterson, B. K., Pennings, P. S., Jensen, J. D. and Hoekstra, H. E. (2012). Evidence of adaptation from ancestral variation in young populations of beach mice. *Evolution* 66, 3209–3223.

Eickwort, G. C., Eickwort, J. M., Gordon, J., Eickwort, M. A. and Wcislo, W. T. (1996). Solitary behavior in a high-altitude population of the social sweat bee *Halictus rubicundus* (Hymenoptera: Halictidae). *Behav. Ecol. Sociobiol.* 38, 227.

Elsik, C. G., Worley, K. C., Bennett, A. K., Beye, M., Camara, F., Childers, C. P., de Graaf, D. C., Debyser, G., Deng, J., Devreese, B., et al. (2014). Finding the missing honey bee genes: Lessons learned from a genome upgrade. *BMC Genomics* 15, 86.

Fahrbach, S. E., Robinson, G. E., Capaldi, E. A., Moore, D. and Farris, S. M. (1998). Experience-expectant plasticity in the mushroom bodies of the honeybee. *Learn. Mem.* 5, 115–123.

Ferreira, P. G., Patalano, S., Chauhan, R., Ffrench-Constant, R., Gabaldón, T., Guigó, R. and Sumner, S. (2013). Transcriptome analyses of primitively eusocial wasps reveal novel insights into the evolution of sociality and the origin of alternative phenotypes. *Genome Biol.* 14, R20.

Gerschenfeld, H. M. (1973). Chemical transmission in invertebrate central nervous systems and neuromuscular junctions. *Physiol. Rev.* 53, 1–119.

Goodisman, M. A. D., Isoe, J., Wheeler, D. E. and Wells, M. A. (2005). Evolution of insect metamorphosis: A microarray-based study of larval and adult gene expression in the ant *Camponotus festinatus*. *Evolution* 59, 858–870.

Greiner, B., Ribi, W. A., Wcislo, W. T. and Warrant, E. J. (2004). Neural organisation in the first optic ganglion of the nocturnal bee *Megalopta genalis*. *Cell Tissue Res.* 318, 429–437.

Hines, H. M., Hunt, J. H., O'Connor, T. K., Gillespie, J. J. and Cameron, S. A. (2007). Multigene phylogeny reveals eusociality evolved twice in vespid wasps. *Proc. Natl. Acad. Sci.* 104, 3295–3299.

Honeybee Genome Sequencing Consortium (2006). Insights into social insects from the genome of the honeybee *Apis mellifera*. *Nature* 443, 931.

Humann, F. C. and Hartfelder, K. (2011). Representational Difference Analysis (RDA) reveals differential expression of conserved as well as novel genes during caste-specific development of the honey bee (*Apis mellifera* L.) ovary. *Insect Biochem. Mol. Biol.* 41, 602–612.

Ismail, N., Robinson, G. E. and Fahrbach, S. E. (2006). Stimulation of muscarinic receptors mimics experience-dependent plasticity in the honey bee brain. *Proc. Natl. Acad. Sci.* 103, 207–211.

Ismail, N., Christine, S., Robinson, G. E. and Fahrbach, S. E. (2008). Pilocarpine improves recognition of nestmates in young honey bees. *Neurosci. Lett.* 439, 178–181.

Jones, B. M., Leonard, A. S., Papaj, D. R. and Gronenberg, W. (2013). Plasticity of the worker bumblebee brain in relation to age and rearing environment. *Brain. Behav. Evol.* 82, 250–261.

Kapheim, K. M., Bernal, S. P., Smith, A. R., Nonacs, P. and Wcislo, W. T. (2011). Support for maternal manipulation of developmental nutrition in a facultatively eusocial bee, *Megalopta genalis* (Halictidae). *Behav. Ecol. Sociobiol.* 65, 1179–1190.

Kapheim, K. M., Wcislo, W. T., Amdam, G. V, Smith, A. R., Ihle, K. E. and Nonacs, P. (2012). Physiological variation as a mechanism for developmental caste-biasing in a facultatively eusocial sweat bee. *Proc. R. Soc. B Biol. Sci.* 279, 1437–1446.

Kapheim, K. M., Smith, A. R., Nonacs, P., Wcislo, W. T. and Wayne, R. K. (2013). Foundress polyphenism and the origins of eusociality in a facultatively eusocial sweat bee, *Megalopta genalis* (Halictidae). *Behav. Ecol. Sociobiol.* 67, 331–340.

Kapheim, K. M., Pan, H., Li, C., Salzberg, S. L., Puiu, D., Magoc, T., Robertson, H.M., Hudson, M.E., Venkat, A., Fischman, B.J. et al. (2015). Genomic signatures of evolutionary transitions from solitary to group living. *Science* 348, 1139–1143.

Kocher, S. D. and Paxton, R. J. (2014). Comparative methods offer powerful insights into social evolution in bees. *Apidologie* 45, 289–305.

Kocher, S., Li, C., Yang, W., Tan, H., Yi, S., Yang, X., Hoekstra, H., Zhang, G., Pierce, N. and Yu, D. (2013). The draft genome of a socially polymorphic halictid bee, *Lasioglossum albipes*. *Genome Biol.* 16, 34.

Martin, A., Papa, R., Nadeau, N. J., Hill, R. I., Counterman, B. A., Halder, G., Jiggins, C. D., Kronforst, M. R., Long, A. D., McMillan, W. O., et al. (2012). Diversification of complex butterfly wing patterns by repeated regulatory evolution of a *Wnt* ligand. *Proc. Natl. Acad. Sci.* 109, 12632–12637.

McCarthy, D. J., Chen, Y. and Smyth, G. K. (2012). Differential expression analysis of multifactor RNA-Seq experiments with respect to biological variation. *Nucleic Acids Res.* 40, 4288–4297.

Mi, H., Muruganujan, A. and Thomas, P. D. (2013). PANTHER in 2013: modeling the evolution of gene function, and other gene attributes, in the context of phylogenetic trees. *Nucleic Acids Res.* 41, D377-86.

Michener, C. D. (1990). Reproduction and castes in social halictine bees. In *Social Insects: An Evolutionary Approach to Castes and Reproduction*, p. 257. New York: Springer-Verlag.

Narumiya, S. and Morii, N. (1993). *rho* gene products, botulinum C3 exoenzyme and cell adhesion. *Cell. Signal.* 5, 9–19.

Oleksiak, M. F., Churchill, G. A. and Crawford, D. L. (2002). Variation in gene expression within and among natural populations. *Nat. Genet.* 32, 261.

Packer, L. (1990). Solitary and eusocial nests in a population of *Augochlorella striata* (Provancher) (Hymenoptera; Halictidae) at the northern edge of its range. *Behav. Ecol. Sociobiol.* 27, 339–344.

Parker, J. S., Mullins, M., Cheang, M. C. U., Leung, S., Voduc, D., Vickery, T., Davies, S., Fauron, C., He, X., Hu, Z., et al. (2009). Supervised risk predictor of breast cancer based on intrinsic subtypes. *J. Clin. Oncol.* 27, 1160.

Pitman, R. M. (1971). Transmitter substances in insects: A review. *Comp. Gen. Pharmacol.* 2, 347–371.

Rehan, S. M., Berens, A. J. and Toth, A. L. (2014). At the brink of eusociality: transcriptomic correlates of worker behaviour in a small carpenter bee. *BMC Evol. Biol.* 14, 260.

Ridley, A. J. (1995). *Rho*-related proteins: actin cytoskeleton and cell cycle. *Curr. Opin. Genet. Dev.* 5, 24–30.

Robinson, M. D. and Oshlack, A. (2010). A scaling normalization method for differential expression analysis of RNA-seq data. *Genome Biol.* 11, R25.

Robinson, M. D., McCarthy, D. J. and Smyth, G. K. (2010). edgeR: A Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* 26, 139–140.

Rosenwald, A., Wright, G., Chan, W. C., Connors, J. M., Campo, E., Fisher, R. I., Gascoyne, R. D., Muller-Hermelink, H. K., Smeland, E. B., Giltnane, J. M., et al. (2002). The use of molecular profiling to predict survival after chemotherapy for diffuse large-B-cell lymphoma. *N. Engl. J. Med.* 346, 1937–1947.

Sancar, A. (1994). Structure and function of DNA photolyase. *Biochemistry* 33, 2–9.

Shields, G., Dubendorfer, A. and Sang, J. H. (1975). Differentiation in vitro of larval cell types from early embryonic cells of *Drosophila melanogaster*. *J. Embryol. Exp. Morphol.* 33, 159–175.

Smith, A. R., Wcislo, W. T. and O'Donnell, S. (2003). Assured fitness returns favor sociality in a mass-provisioning sweat bee, *Megalopta genalis* (Hymenoptera: Halictidae). *Behav. Ecol. Sociobiol.* 54, 14–21.

Technau, G. and Heisenberg, M. (1982). Neural reorganization during metamorphosis of the corpora pedunculata in *Drosophila melanogaster*. *Nature* 295, 405.

Wcislo, W. T. and Danforth, B. N. (1997). Secondarily solitary: the evolutionary loss of social behavior. *Trends Ecol. Evol.* 12, 468.

Wcislo, W. T., Arneson, L., Roesch, K., Gonzalez, V., Smith, A. and Fernández, H. (2004). The evolution of nocturnal behaviour in sweat bees, *Megalopta genalis* and *M. ecuadoria* (Hymenoptera: Halictidae): an escape from competitors and enemies? *Biol. J. Linn. Soc.* 83, 377–387.

White, K. and Kankel, D. R. (1978). Patterns of cell division and cell movement in the formation of the imaginal nervous system in *Drosophila melanogaster*. *Dev. Biol.* 65, 296–321.

Whitehead, A. (2012). Comparative genomics in ecological physiology: toward a more nuanced understanding of acclimation and adaptation. *J. Exp. Biol.* 215, 884–891.

Whitehead, A. and Crawford, D. L. (2006). Neutral and adaptive variation in gene expression. *Proc. Natl. Acad. Sci.* 103, 5425–5430.

Wolda, H. and Roubik, D. W. (1986). Nocturnal bee abundance and seasonal bee activity in a Panamanian forest. *Ecology* 67, 426–433.

Zayed, A., Naeger, N. L., Rodriguez-Zas, S. L. and Robinson, G. E. (2012). Common and novel transcriptional routes to behavioral maturation in worker and male honey bees. *Genes, Brain Behav.* 11, 253–261.

CHAPTER 4

CASTE-BIASED GENE EXPRESSION IN A FACULTATIVELY EUSOCIAL BEE SUGGESTS A ROLE FOR GENETIC ACCOMMODATION IN THE EVOLUTION OF EUSOCIALITY[1]

Abstract

Developmental plasticity may accelerate the evolution of phenotypic novelty through genetic accommodation, but studies of genetic accommodation often lack knowledge of the ancestral state to place selected traits in an evolutionary context. A promising approach for assessing genetic accommodation involves using a comparative framework to ask whether ancestral plasticity is related to the evolution of a particular trait. Bees are an excellent group for such comparisons because caste-based societies (eusociality) have evolved multiple times independently and extant species exhibit different modes of eusociality. We measured brain and abdominal gene expression in a facultatively eusocial bee, *Megalopta genalis*, and assessed whether plasticity in this species is functionally linked to eusocial traits in other bee lineages. Caste-biased abdominal genes in *M. genalis* overlapped significantly with caste-biased genes in obligately eusocial bees. Moreover, caste-biased genes in *M. genalis* overlapped significantly with genes shown to be rapidly evolving in multiple studies of 10 bee species, particularly for genes in the glycolysis pathway and other genes involved in metabolism. These results provide support for the idea that eusociality can evolve via genetic accommodation, with plasticity in facultatively eusocial species like *M. genalis* providing a substrate for selection during the evolution of caste in obligately eusocial lineages.

---

[1] The work in this chapter has been adapted from a previously published article: Jones, BM, Kingwell, CJ, Wcislo, WT and Robinson, GE. 2017. Caste-biased gene expression in a facultatively eusocial bee suggests a role for genetic accommodation in the evolution of eusociality. *Proc. R. Soc. B* 284 (1846):20162228. Figures and tables have been renumbered. Supplemental materials and files are available online at https://royalsocietypublishing.org/doi/suppl/10.1098/rspb.2016.2228. Reprinted with permission from The Royal Society: https://royalsociety.org/journals/permissions/

Introduction

Phenotypic plasticity may accelerate the evolution of phenotypic novelty through genetic accommodation (Leichty et al., 2012; Pichancourt and van Klinken, 2012; West-Eberhard, 2003a; Yampolsky et al., 2012, c.f. Ghalambor et al., 2015), but studies of genetic accommodation often lack knowledge of the ancestral state to place selected traits in an evolutionary context. Most empirical support for genetic accommodation employs artificial selection on experimentally induced phenotypes (Suzuki and Nijhout, 2006; Waddington, 1942; Waddington, 1953). Other empirical studies demonstrate the necessary phylogenetic relationships between environmentally sensitive phenotypes in ancestral lineages and more fixed phenotypes in derived lineages required for genetic accommodation (Heil et al., 2004; Santana and Dumont, 2009) but do not show evidence of selection. For condition-sensitive traits, the ancestral condition cannot always be inferred accurately from traits of extant forms (Piperno et al., 2015).

As genomic tools are deployed across a wider array of species, a promising approach to investigate genetic accommodation is to look for both plasticity in ancestral lineages and evidence of selection (Ledon-Rettig et al., 2008; Leichty et al., 2012; Moczek et al., 2011). Groups with repeated evolution of traits are ideal for this approach, as comparative genomics can be used to identify genes under selection for a trait of interest. Social insects, and bees in particular, are especially promising for addressing the role of genetic accommodation in phenotypic innovation because eusocial behavior has evolved independently multiple times (Brady et al., 2006; Cameron and Mardulyn, 2001; Cardinal et al., 2010; Schwarz et al., 2007).

A growing volume of published data has identified genes undergoing selection in social lineages of bees (Harpur et al., 2014; Kapheim et al., 2015; Woodard et al., 2011), providing a framework to test the role of genetic accommodation in the evolution of eusociality. Still missing, however, is knowledge of the ancestral phenotypic plasticity of different bee lineages, and whether this ancestral plasticity is related to genes under selection in eusociality.

Gene expression data have been used as a measure of plasticity in many different contexts, including those related to genetic accommodation (Leichty et al., 2012). In particular, genes that are differentially expressed are thought to experience reduced genetic constraint and evolve more quickly, which has been confirmed in empirical studies (Hunt et al., 2010; Snell-Rood et al., 2011). Brain gene expression differences have been described across caste and many

behavioral contexts in bumblebees (Harrison et al., 2015) and honeybees (Ament et al., 2012; Chandrasekaran et al., 2011). Queen (Q) and worker (W) honeybees are morphologically distinct and adapted to reproductive and non-reproductive functions, respectively; these differences are reflected in a transcriptomic study which reported over 2000 differentially expressed genes (DEGs) in the brain (Grozinger et al., 2007). These genes, along with caste-biased genes in other obligately eusocial insects, have been found to evolve more rapidly than genes unrelated to caste (Harpur et al., 2014; Hunt et al., 2010; Hunt et al., 2011). However, whether these genes also showed facultative expression along with plasticity in organismal-level phenotypic traits prior to the evolution of obligate eusociality is unknown.

Robust phylogenetic studies point to a solitary ancestral lifestyle for bees (Wilson, 1971; Wilson and Holldobler, 2005). Some species of bees display facultative eusociality, with both solitary and social nests existing either across geographical gradients (Cronin and Hirata, 2003; Eickwort et al., 1996; Field et al., 2010; Packer, 1990; Plateaux-Quénu et al., 2000, reviewed in Kocher and Paxton, 2014) or even within the same population (Wcislo et al., 2004; Yanega, 1988). It is thus likely that at least some mechanisms underpinning phenotypic differentiation in these facultatively eusocial bees play roles in evolutionary transitions from solitary to social life histories. If so, then testing the predictions of genetic accommodation in eusocial evolution can use facultatively eusocial species as proxies for the ancestral state. We did this by using the facultatively eusocial bee, *Megalopta genalis,* to measure environmentally induced plasticity in gene expression and then compared differentially expressed genes from this species with genes previously shown to be under selection in obligately eusocial taxa.

Megalopta genalis (Halictidae) is a Neotropical sweat bee that displays facultative eusociality (Kapheim et al., 2011; Kapheim et al., 2012; Wcislo and Gonzalez, 2006; Wcislo et al., 2004). Many facultatively eusocial species exhibit social plasticity across geographical gradients (Cronin and Hirata, 2003; Eickwort et al., 1996; Plateaux-Quénu et al., 2000), but in *M. genalis* both social and solitary nests exist within a single population. This strongly suggests that *M. genalis* eusociality is at least partially environmentally determined. The plasticity arises through variation in reproductive behavior of nest-founding females (Kapheim et al., 2011; Kapheim et al., 2012; Kapheim et al., 2013; Smith et al., 2009). Solitary nests form when females produce only males in their first broods, with subsequent female production resulting in dispersal rather than retention of female workers (Kapheim et al., 2013). By contrast, social nests

form when one or more females produced in the first brood remain as non-reproductive workers. Expression of alternative reproductive phenotypes in *Megalopta* is related to social competition linked to body size and nutrition (Kapheim et al., 2011; Kapheim et al., 2013; Smith et al., 2008), as well as ovary size (Kapheim et al., 2012) and hormonal differences (Smith et al., 2013). Workers in social nests remain sensitive to environmental and social conditions and have the ability to mate and reproduce after queen loss or supersedure, becoming replacement queens with reproductive outputs equivalent to those of solitary females (Smith et al., 2009).

We used the naturally occurring phenotypic variation of *M. genalis* and comparative genomics to explore the mechanisms of genetic accommodation. We compared gene expression for four female phenotypes (solitary, queen, worker and replacement queen) in both brain and abdominal tissues. We then assessed the degree to which differences in expression associated with phenotypic variation in *M. genalis* are common across other species of bees to ask whether similar molecular mechanisms are implicated in caste determination. Finally, we tested for commonality of caste-biased genes in *M. genalis* with genes previously identified as undergoing selection in eusocial bee lineages to address whether ancestral plasticity is consistent with genetic accommodation for social traits in the evolution of eusociality in bees.

Material and Methods

Sample collection

Megalopta genalis females were collected on Barro Colorado Island in the Republic of Panama where they are abundant during the dry season (Wolda and Roubik, 1986). Frozen tissues were exported with permission from the National Authority for the Environment of the Government of Panama (permit Nos. SEX/A-53-13 and SEX/AH-4-15). Observation nests were created with newly emerged females and monitored daily until offspring emergence enabled classification of nests as solitary or social (Kapheim et al., 2011; Smith et al., 2007; Wcislo and Gonzalez, 2006; Wcislo et al., 2004). Solitary females (S) and queens (Q) were collected following an egg-laying event in solitary and social nests, respectively. In half of the social nests, workers (W) were removed on the same day as queens, and in the other half workers were allowed to transition into replacement queens (R) before collection (4–18 days following queen removal). All females were collected during inactive periods so that gene expression differences were not likely due to acute differences in activity or short-term behaviors, but rather stable gene

expression differences between groups. Detailed sample information is provided in the electronic supplementary material, S1.

RNA preparation

Whole brains (n = 30; 7 S, 7 Q, 9 W, 7 R) and abdomens (n = 25; 7 S, 7 Q, 6 W, 5 R; gut tissue and Dufour's glands removed) were dissected as in Jones et al. (2015), and ovaries were imaged to confirm reproductive state. Abdominal tissues extracted were primarily fat body and ovarian tissue, but also include the sting sac, muscle and nervous tissues. To facilitate communication, results will refer to 'abdominal tissues' which denotes these tissues collectively. Total RNA was extracted using QIAGEN RNeasy columns following the manufacturer's protocol.

Library preparation and RNA-sequencing

Poly-A RNA was enriched from total RNA and strand-specific cDNA libraries were prepared using the Bioo Scientific NEXTflex Directional RNA-Sequencing Kit (dUTP Based) for Illumina. Paired-end sequencing was performed on an Illumina HiSeq 2500 at the W. M. Keck Center (University of Illinois). Quality was assessed using FASTQC and read trimming was performed with Trimmomatic prior to alignment with Bowtie to a previously assembled transcriptome for *M. genalis* (Jones et al., 2015).

Differential expression analyses

Estimated read counts at the putative gene level were obtained using RSEM following alignment with scripts packaged in Trinity r20140413 (Grabherr et al., 2011). Gene counts were filtered to include only genes with at least 1 count per million in the minimum number of samples per group per tissue type. A surrogate variable analysis (Leek et al., 2010; Leek et al., 2012) was performed on each tissue dataset to identify potential batch effects due to collection year, library preparation batch, sequencing lane or other unidentified technical differences. Dispersion estimates and pairwise tests of differential expression were conducted in edgeR (McCarthy et al., 2012) with FDR correction. DEG lists can be found in the electronic supplementary material, S2. Annotation of differentially expressed genes for GO enrichment was conducted using PANTHER for pairwise lists of DEGs, and a statistical overrepresentation test with Bonferroni correction for multiple testing was used to identify PANTHER pathways and GO-Slim Biological Processes overrepresented in pairwise lists as presented in the electronic supplementary material, S3 (Mi et al., 2013).

Overlap with other studies

Putative orthologues between species were identified using BLAST reciprocal best hits (RBH) between predicted peptides. For *M. genalis* and *Bombus terrestris* transcriptomes (Colgan et al., 2011; Jones et al., 2015), predicted peptides were obtained using TransDecoder. Conversion lists between microarray probes and annotation versions of the *Apis mellifera* genome, along with RBH results, are found in the electronic supplementary material, S4. For Representation Factor (RF) gene overlap tests, only genes tested in the study with a putative orthologue in both species were compared between any two given studies. Gene lists and complete RF results are given in the electronic supplementary material, S5 (gene expression studies) and S6 (selection studies).

Results

Caste differences in gene expression

Female castes of *M. genalis* differed in gene expression in both brain and abdominal tissues (Fig. 4.1 and Table 4.1). In abdominal tissues, variance in gene expression was largely explained by reproductive activity. The first principal component (Fig. 4.1A) explained nearly 40% of the variance in abdominal gene expression and separated workers (who are non-reproductive) from all reproductive groups. All reproductive females showed similar patterns of abdominal gene expression relative to workers, regardless of sociality (Fig. 4.2). Nearly 95% (3618/3827) of the genes that were more highly expressed in solitary females compared to workers were also more highly expressed in queens compared with workers, a highly significant overlap (representation factor: 3.34, $p < 0.0001$). These 3618 genes were strongly enriched for GO-Slim Biological Processes related to DNA metabolism and repair, chromatin organization and cell cycle (all GO enrichments listed in the electronic supplementary material, S3).

Among abdominal worker-biased genes (Fig. 4.2A), 2629 genes were more highly expressed in both W > Q and W > S comparisons. The most enriched GO-Slim Biological Process for these genes was glycolysis (GO:0006096), with a nearly fivefold enrichment. Also enriched in worker-biased abdominal genes compared with queens and solitary females were steroid metabolic process (GO:0008202), respiratory electron transport chain (GO:0022904) and monosaccharide metabolic process (GO:0005996).

Brain differences in expression were much less pronounced than abdominal differences

(Table 4.1), but variance in brain gene expression explained by the first two principal components (22.6% of total variance) roughly mapped to the variance in behavior seen across castes. Queens and workers were least variable, while both solitary and replacement queen females showed large variation in brain gene expression. The latter are the two groups that, at the time of collection, performed all activities in the nest (cell building, foraging, provisioning and reproduction). Many of the DEG lists in brain tissue did not have significant functional enrichment. However, heterotrimeric G-protein signalling pathways (P00026 and P00027) and both muscarinic (P00042 and P00043) and nicotinic (P00044) acetylcholine receptor signalling pathways were upregulated in both workers and replacement queens relative to solitary females. In solitary females, extracellular transport (GO:0006858) was enriched relative to workers.

Replacement queens shift gene expression to reproductive-like phenotypes

The shift to reproductive activity in replacement queens following queen removal was associated with a near-complete shift in abdominal gene expression from worker-like to queen- and solitary-like (Fig. 4.1A and Fig. 4.2). Of the 3618 genes in common between Q > W and S > W abdominal DEGs, 81% were also more highly expressed in replacement queens compared with workers.

Replacement queens displayed high variation in brain gene expression compared with workers and queens, possibly reflecting the increased behavioral variation in this group. Three individuals (all collected in 2014) more closely resembled the reproductive phenotype along the first two principal components shown in Fig. 4.1B, while four individuals (two collected in 2014 and two in 2015) appeared worker-like in brain gene expression.

Cross-species expression comparisons

Abdominal caste-specific differences in gene expression in *M. genalis* overlapped significantly with caste-specific DEGs in *B. terrestris* and five worker-related DEG lists in *A. mellifera.* In addition, brain caste-specific differences in gene expression in *M. genalis* overlapped significantly with differences in gene expression in two out of four comparisons with *A. mellifera* DEGs.

Caste-specific abdominal expression in *M. genalis* was compared with previous studies of *B. terrestris* caste differences (Harrison et al., 2015) and *A. mellifera* workers (Ament et al., 2011; Galbraith et al., 2016). Significant overlap was observed when comparing queen-biased and worker-biased genes in *M. genalis* abdomen and *B. terrestris* whole-body extractions (Fig.

4.3; RF: 1.5 for Q > W and RF: 1.6 for W > Q, both $p < 0.0001$). Of the genes more highly expressed in *B. terrestris* queens, 49% were also more highly expressed in the abdomens of *M. genalis* queens compared with workers. These genes were enriched for DNA metabolic process (GO:0006259), cell cycle (GO:0007049) and nucleobase-containing compound metabolic process (GO:0006139).

Very strong overlap was observed when comparing queen-biased and worker-biased genes in *M. genalis* abdomens with DEGs in the abdomens of laying worker (LW) and sterile worker (S) honeybees (Fig. 4.3; RF: 2.0 for Q > W,LW > S and 1.7 for W > Q,S > LW; both $p < 0.0001$; Galbraith et al., 2016). Among the reproductive-related (Q > W and LW > S) overlap, enriched GO-Slim terms included chromatin organization (GO:0006325), DNA replication (GO:0006260) and mitosis (GO:0007067). No significant GO enrichment was observed for the overlapping genes between *M. genalis* W > Q and *A. mellifera* S > LW.

Significant overlap was also observed when comparing the *M. genalis* Q versus W abdomen DEG list and genes differentially expressed in the fat body of worker honeybees fed high versus low pollen diets (Ament et al., 2011): Q > W genes had marginal overlap with Low > High pollen (RF: 1.1, $p < 0.049$) and W > Q genes had strong overlap with High > Low pollen-responsive genes (RF: 1.6, $p< 0.0001$; Fig. 4.3). For the remaining comparisons, *M. genalis* workers showed gene expression consistent with both worker phenotypes in honeybees. Genes that were more highly expressed in the abdomens of workers than in queens of *M. genalis* overlapped significantly with nurse (N) and forager (F) biased genes, genes both upregulated and downregulated by exposure to QMP, and vitellogenin (*Vg*)-responsive genes in the honeybee fat body (Ament et al., 2011, Fig. 4.3).

No significant overlap was observed when comparing brain Q versus W DEGs in *M. genalis* and *A. mellifera* (Grozinger et al., 2007). The majority of *A. mellifera* brain gene expression studies have compared behaviorally distinct subcastes of workers. Honeybee workers as a whole resemble *M. genalis* castes in terms of their behavioral flexibility (e.g. performance of both nurse-like and forager-like tasks). We therefore tested for overlap of our *M. genalis* Q versus W brain DEGs and DEGs from two *A. mellifera* N versus F experiments (Alaux et al., 2009; Whitfield et al., 2003), as well as from an RNAi experiment looking at the effects of peripheral *Vg* knockdown on worker brain gene expression (Wheeler et al., 2013).

Significant overlap (RF: 2.0, $p < 0.021$) was found when comparing *M. genalis* Q versus

W DEGs and DEGs from one N versus F study (Alaux et al., 2009) but not the other ($p < 0.379$, Whitfield et al., 2003). Among the 11 overlapping DEGs between the *M. genalis* Q versus W and *A. mellifera* N versus F lists are *serine/threonine-protein kinase ICK-like, apyrase precursor* and two transporter proteins. A significant degree of overlap was found when comparing worker-biased genes in *M. genalis* (W > Q) and genes more highly expressed in *Vg* knockdown bees relative to control bees (Wheeler et al., 2013) (RF: 2.1, $p < 0.017$). Four of the 10 *Vg*RNAi-overlapping genes were also within the 11 genes that overlap between *M. genalis* Q versus W and *A. mellifera* N versus F lists, including *serine/threonine-protein kinase ICK-like, apyrase precursor,* and an excitatory amino acid transporter. Lists of overlapping genes and tests for significance are found in the electronic supplementary material, S5.

Insights from comparative molecular evolution studies

Worker-biased genes expressed in the abdomen of *M. genalis* were enriched for many GO terms related to metabolism. This was noteworthy because previous molecular evolution studies identified metabolic genes as one of the more prominent categories of genes undergoing selection in social bee lineages (Fig. 4.4; Kapheim et al., 2015; Woodard et al., 2011). To more formally test the association between worker-biased genes in *M. genalis* and genes under selection in other bee species, we compared our abdomen W > Q DEG list with genes identified in three studies of molecular evolution (Harpur et al., 2014; Kapheim et al., 2015; Woodard et al., 2011). In all three comparisons, the genes that were more highly expressed in the abdomens of workers relative to queens in *M. genalis* were also overrepresented among genes undergoing positive selection in social lineages of bees (Table 4.2; electronic supplementary material, S6).

Among the Harpur *et al.* (2014) genes showing positive selection in *A. mellifera* that have putative *M. genalis* orthologues, nearly 40% (254/639) were worker-biased in expression (RF: 1.4, $p < 0.0001$; Table 4.2). These genes were enriched for GO processes relating to respiratory electron transport chain (GO:0022904) and the generation of precursor metabolites and energy (GO:0006091). In contrast with worker-biased genes, queen-biased genes were underrepresented among the Harpur *et al.* (2014) genes showing positive selection (RF: 0.7, $p < 0.0001$). Genes identified under positive selection in Woodard *et al.* (2011) included 50 that were also worker-biased in expression in *M. genalis*, eight of which are in the glycolysis pathway (Fig. 4.4; RF: 1.3, $p = 0.015$). These 50 genes were enriched for cellular amino acid metabolic process (GO:0006520), and seven were also among the Harpur *et al.* (2014) genes, including *aldehyde*

oxidase-like, aspartate aminotransferase, prostaglandin reductase 1-like, bifunctional ATP-dependent dihydroxyacetone kinase and 6-phosphofructokinase.

Genes identified as undergoing positive selection (103, 70 of which have *M. genalis* putative orthologues) across social lineages in Kapheim *et al.* (2015) also overlapped significantly with abdominal worker-biased genes in *M. genalis* (28 genes; RF: 1.5, $p = 0.014$). These genes were not enriched for GO terms, but four of the 28 genes are in the glycolysis pathway (Fig. 4.4), including *beta-lactamase-like protein 2 homologue, glyceraldehyde-3-phosphate dehydrogenase 2 isoform 1, phosphoglycerate mutase 2-like* and *enolase-like.*

Discussion

The fact that eusociality has evolved independently multiple times in bees has provided rich material for comparative analyses. Recent work has demonstrated that different origins of eusociality have involved both common and unique pathways, often with changes in gene regulatory networks (Berens et al., 2014; Kapheim et al., 2015; Woodard et al., 2011). We found that abdominal caste-specific differences in gene expression in a facultatively eusocial bee, *M. genalis,* overlapped significantly with caste-specific differences in gene expression in other eusocial bee species. Our findings suggest common mechanisms involved in caste regulation across lineages with different types of eusociality. Moreover, we found that worker-biased abdominal genes in *M. genalis* overlapped significantly with genes shown to be rapidly evolving in eusocial lineages of bees from three independent studies of selection. With *M. genalis* serving as a proxy for ancestral variation, these results provide support for genetic accommodation in the evolution of eusociality in bees. According to this hypothesis, phenotypic plasticity exhibited by facultatively eusocial species may have enabled selection for permanent reproductive and non-reproductive castes.

Dramatic differences in gene expression were observed in abdominal tissues when comparing workers with reproductively active females. While DNA-related processes dominated the reproductive signal in abdominal tissues, workers showed a bias for expression of many metabolic pathways, including glycolysis (Fig. 4.4). In contrast with the abdominal results, brain gene expression differences were more subtle. The number of DEGs between queens and workers was more similar to that seen in the primitively eusocial wasp *Polistes dominula* (Standage et al., 2016) than in the highly derived eusocial honeybee (Grozinger et al., 2007). The

striking differences in DEG numbers between the brain and abdomen suggest that signalling between the reproductive system and the brain plays a role in mediating behavior of *M. genalis* females. This is similar to hypotheses from the honeybee literature (Amdam et al., 2006; Wang et al., 2009; Wang et al., 2010), where it has been reported that surgical implantation of supernumerary ovaries in workers leads to a shift in the age at onset of foraging (Wang et al., 2010). An alternative explanation for the differences in DEG numbers between brain and abdominal comparisons is that collection of individuals during inactive periods led to a reduction in the genes that were differentially expressed in the brain.

Theory predicts that social traits may be subject to relaxed selective constraint and higher levels of polyphenism (Gadagkar, 1997; Linksvayer and Wade, 2009). Our results are consistent with predicted relaxed constraints on worker traits, which may predispose worker-biased genes to accumulate mutations to be screened by selection (Moczek et al., 2011; Van Dyken and Wade, 2010). Consistent with theory, the genes identified by Harpur *et al*. (2014) with signatures of adaptive evolution in honeybees were also worker-biased, and worker-biased genes have been reported to be more derived in multiple social insect species (Feldmeyer et al., 2014; Ferreira et al., 2013; Harpur et al., 2014). Overlap of *M. genalis* worker-biased genes and those under selection in eusocial lineages include many involved in glycolysis (Fig. 4.4), which has been implicated in caste determination not only in bees but also in other social insects (Berens et al., 2014).

Phenotypes that are originally environmentally induced can be selected upon and shaped, such that inherited variants can express the trait absent the environmental induction (West-Eberhard, 2003). Genetic accommodation predicts selection on genes that are environmentally sensitive if there is a selective advantage to the environmentally induced phenotype. Eusocial insects have evolved a worker caste that is reproductively inactive and often specialized for non-reproductive behaviors relative to the ancestral solitary state. In multiple social insect lineages, this specialization was enabled by selection on genes that are now worker-biased in their expression (Feldmeyer et al., 2014; Ferreira et al., 2013; Harpur et al., 2014). Our study implicates some of the same genes in the flexible worker phenotype of *M. genalis,* which suggests a role for genetic accommodation in the evolution of specialized worker castes. If genetic accommodation facilitated worker specialization, we would expect ancestral plasticity in gene expression associated with worker-related genes. Consistent with this prediction, we found

that worker-biased genes in *M. genalis* share significant identity with genes identified as undergoing selection in three independent tests of positive selection across eusocial lineages of bees (Harpur et al., 2014; Kapheim et al., 2015; Woodard et al., 2011). Plasticity in these genes in the ancestors of obligately eusocial species may have facilitated the evolution of derived worker castes through genetic accommodation, leading to a more fixed caste determination system as seen in many obligately eusocial groups.

We provide support for the idea that genetic accommodation may have played a role in the evolution of caste, with ancestral environmentally induced plasticity leading to selection on worker traits in the evolution of eusociality in bees. Future research investigating allele frequency change within *M. genalis* itself would strengthen this support if caste-biased genes were found to be under selection in this species. As molecular data from other incipiently social lineages are integrated with knowledge from complex eusocial species (Kocher and Paxton, 2014; Rehan and Toth, 2015), it will be possible to explore the role of genetic accommodation in eusociality more rigorously.

Data accessibility

Raw sequencing data have been deposited in the short read archive (NCBI) under the accession SRP079663. The transcriptome used for alignment is available as a TSA submission under the accession PRJNA282469. Electronic supplementary material associated with this chapter is available online.

Acknowledgments

I would like to thank Callum Kingwell and William Wcislo for their contributions as co-authors. Betzi Perez, Gabriel Trujillo, Isis Lopez, Esther Velasquez and Ernesto Gomez-Perez assisted with collections in the field. I also thank Amy Cash Ahmed for assistance with molecular work, and Claudia Lutz and members of the Robinson lab for helpful discussion and comments on the text.

Figures and Tables

Figure 4.1

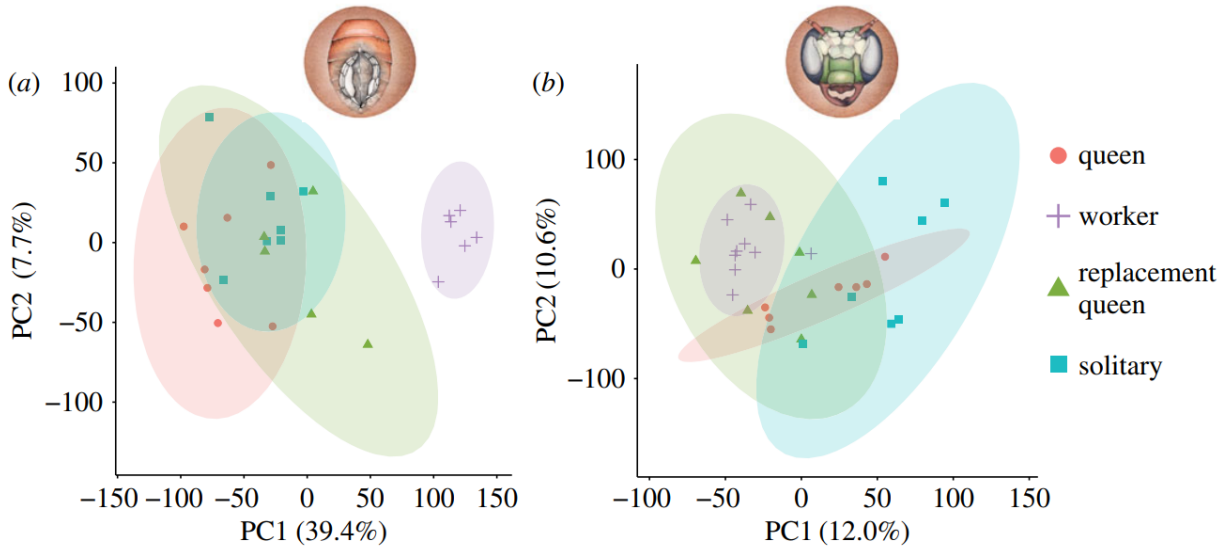


Figure 4.1. Plot of first two principal components for (a) abdominal and (b) brain gene expression patterns across four female behavioral groups. Points represent individual samples and shaded ellipses show 95% CIs. Percentage of variance explained by each principal component is shown on axes. Drawings of *M. genalis* brain and abdominal tissues by Julie Himes.

Figure 4.2

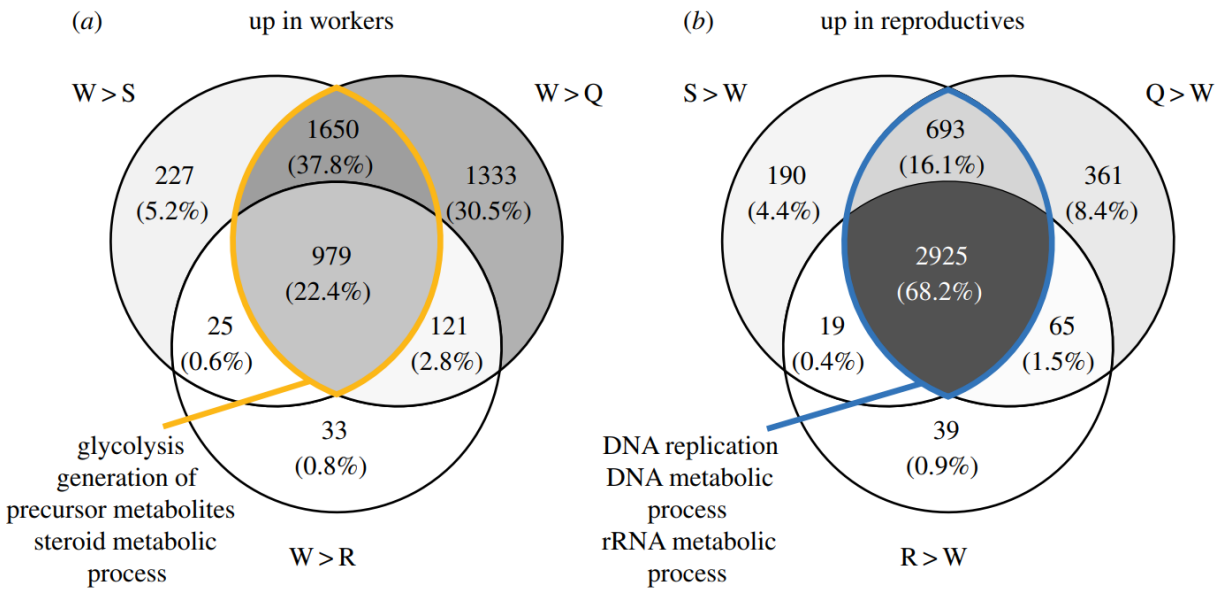


Figure 4.2. Venn diagrams show genes (a) upregulated in workers relative to reproductive individuals and (b) upregulated in reproductive individuals relative to workers. Numbers in each subregion show the number of DEGs with an FDR < 0.05, and percentage of genes in that subregion relative to the universe of genes. Shading is relative to percentage of genes in each subregion, with darker regions corresponding to regions with higher percentages of genes. Yellow and blue highlighted regions show overlap between solitary and queen versus worker DEGs, respectively, and text refers to the top 3 significant enriched GO terms for genes in those regions. Q, queen; W, worker; S, solitary reproductive; R, replacement queen.

Figure 4.3

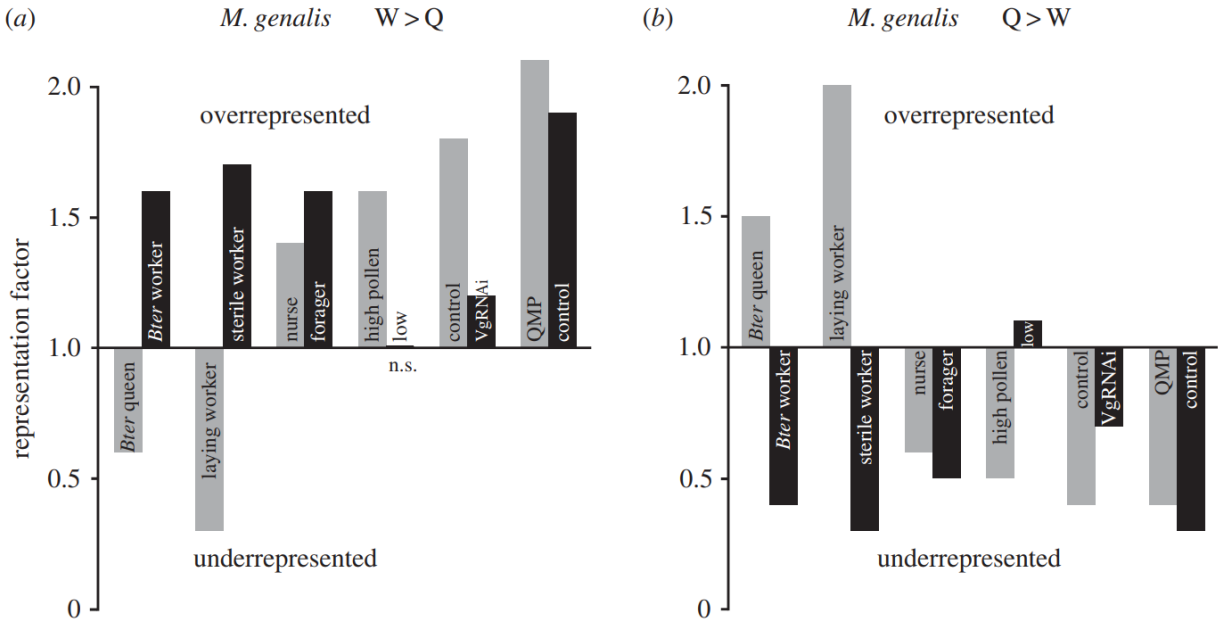
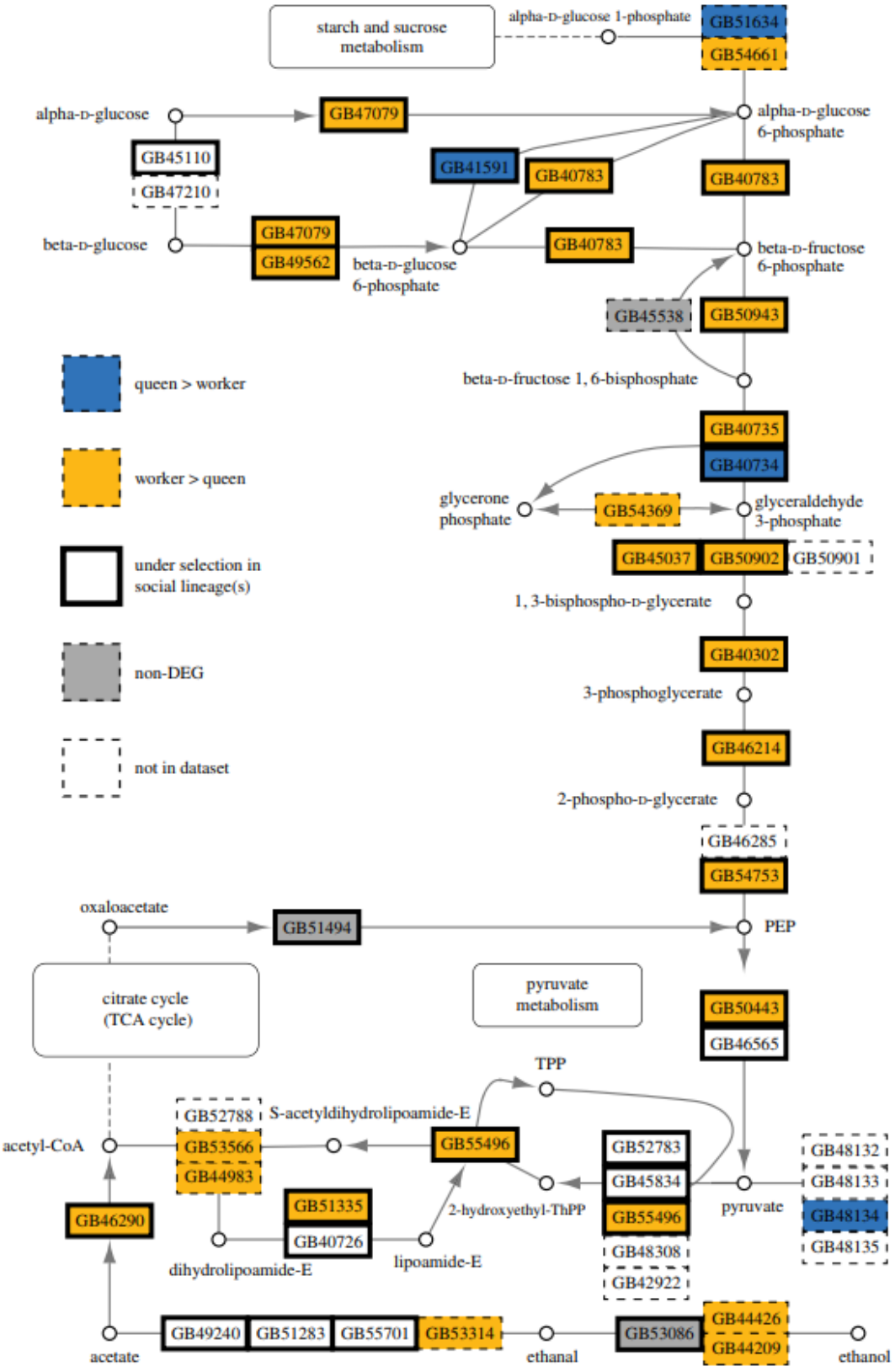


Figure 4.3. Representation factors for overlap of *M. genalis* abdominal (a) worker > queen DEGs and (b) queen > worker DEGs with previous studies in *Bombus terrestris* (whole body, Harrison et al., 2015) and *Apis mellifera* (laying versus sterile worker, abdomen, Galbraith et al., 2016; all other comparisons, fat body, Ament et al., 2011). For each comparison, DEGs are split into two bars and labelled with the group more highly expressed (e.g. the black bar labelled '*Bter* worker' in (a) shows the representation factor for W > Q genes in *M. genalis* and W > Q genes in *B. terrestris*). A representation factor (RF) of 1 indicates a level of overlap expected by chance, while RF > 1 indicates more overlap than expected and RF < 1 indicates less overlap than expected; n.s. not significantly different from RF = 1 (hypergeometric $p > 0.05$), all other bars are significantly different from RF = 1 (hypergeometric $p < 0.05$) in the direction shown.

Figure 4.4


starch and sucrose metabolism
alpha-D-glucose 1-phosphate
GB51634
GB54661
alpha-D-glucose
GB47079
alpha-D-glucose 6-phosphate
GB45110
GB47210
GB41591
GB40783
GB40783
beta-D-glucose
GB47079
GB49562
beta-D-glucose 6-phosphate
GB40783
beta-D-fructose 6-phosphate
GB45538
GB50943
beta-D-fructose 1, 6-bisphosphate
GB40735
GB40734
glycerone phosphate
GB54369
glyceraldehyde 3-phosphate
GB45037
GB50902
GB50901
1, 3-bisphospho-D-glycerate
GB40302
3-phosphoglycerate
GB46214
2-phospho-D-glycerate
GB46285
GB54753
PEP
queen > worker
worker > queen
under selection in social lineage(s)
non-DEG
not in dataset
oxaloacetate
GB51494
citrate cycle (TCA cycle)
pyruvate metabolism
GB50443
GB46565
TPP
GB52788
S-acetyldihydrolipoamide-E
acetyl-CoA
GB53566
GB44983
GB55496
GB52783
GB45834
GB55496
GB48308
GB42922
GB48132
GB48133
GB48134
GB48135
pyruvate
GB46290
GB51335
GB40726
dihydrolipoamide-E
lipoamide-E
2-hydroxyethyl-ThPP
acetate
GB49240
GB51283
GB55701
GB53314
ethanal
GB53086
GB44426
GB44209
ethanol

(Figure 4.4 caption continued)

Figure 4.4. Differentially expressed genes (FDR < 0.05) between queen and worker *M. genalis* abdominal tissues in the glycolysis/gluconeogenesis pathway. DEGs were mapped onto putative honeybee orthologues modified from KEGG pathway ame00010 (Kyoto Encyclopedia of Genes and Genomes, http://www.genome.jp/kegg/) and Woodard et al. (2011). Genes labelled as under selection were identified as undergoing positive selection in at least one of three studies of selection in bees (Harpur et al., 2014; Kapheim et al., 2015; Woodard et al., 2011). Genes not in dataset are honeybee genes without a BLAST reciprocal best hit to the *M. genalis* transcriptome or genes that are not expressed in the abdomen of *M. genalis*.

Table 4.1. Number of differentially expressed genes (DEGs, FDR , 0.05) for each pairwise comparison of female groups in both brain and abdominal tissues. Numbers in parentheses indicate number of genes more highly expressed in first group of pair (e.g. Q versus W comparison has 157 total DEGs, 49 of which are more highly expressed in Q compared to W). Q, queen; W, worker; S, solitary reproductive; R, replacement queen.

Comparison	Abdomen DEGs	Brain DEGs
Q versus W	8127 (4044)	157 (49)
S versus W	6708 (3827)	542 (219)
R versus W	4206 (3048)	0
Q versus R	510 (88)	18 (8)
S versus R	38 (9)	133 (30)
S versus Q	37 (35)	16 (11)

Table 4.2. Overlap of genes more highly expressed in the abdomens of *M. genalis* workers than queens and three independent studies of selection across bees. Woodard et al. (2011) genes are those identified as rapidly evolving in highly eusocial lineages of bees. Harpur et al. (2014) genes are those with signatures of positive selection in *A. mellifera*. Kapheim et al. (2015) genes are those undergoing positive selection across two independent origins of eusociality. Gene lists were restricted to those that had a putative orthologue in *M. genalis* (based on BLAST reciprocal best hit) and were expressed in the abdomen of *M. genalis* females. RF: Representation Factor, *p*-value is from hypergeometric test of overlap. GO terms listed are from PANTHER overrepresentation tests (Bonferroni corrected $p < 0.05$).

Source	**Genes under selection**	**No. W>Q**	**RF**	***p*-value**	**GO enrichment**
Woodard et al.	130	50	1.3	0.015	cellular amino acid metabolic process
Harpur et al.	639	254	1.4	<0.0001	respiratory electron transport chain, generation of precursor metabolites and energy
Kapheim et al.	70	28	1.5	0.014	none

References

Alaux, C., Le Conte, Y., Adams, H. A., Rodriguez-Zas, S., Grozinger, C. M., Sinha, S. and Robinson, G. E. (2009). Regulation of brain gene expression in honey bees by brood pheromone. *Genes, Brain Behav.* 8, 309.

Amdam, G. V, Csondes, A., Fondrk, M. K. and Page Jr, R. E. (2006). Complex social behaviour derived from maternal reproductive traits. *Nature* 439, 76-78.

Ament, S. A., Chan, Q. W., Wheeler, M. M., Nixon, S. E., Johnson, S. P., Rodriguez-Zas, S. L., Foster, L. J. and Robinson, G. E. (2011). Mechanisms of stable lipid loss in a social insect. *J. Exp. Biol.* 214, 3808–3821.

Ament, S. A., Blatti, C. A., Alaux, C., Wheeler, M. M., Toth, A. L., Le Conte, Y., Hunt, G. J., Guzman-Novoa, E., DeGrandi-Hoffman, G., Uribe-Rubio, J. L., et al. (2012). New meta-analysis tools reveal common transcriptional regulatory basis for multiple determinants of behavior. *Proc. Natl. Acad. Sci.* 109, E1801–E1810.

Berens, A. J., Hunt, J. H. and Toth, A. L. (2014). Comparative transcriptomics of convergent evolution: Different genes but conserved pathways underlie caste phenotypes across lineages of eusocial insects. *Mol. Biol. Evol.* 32, 690-703.

Brady, S. G., Sipes, S., Pearson, A. and Danforth, B. N. (2006). Recent and simultaneous origins of eusociality in halictid bees. *Proc. R. Soc. B Biol. Sci.* 273, 1643–1649.

Cameron, S. A. and Mardulyn, P. (2001). Multiple molecular data sets suggest independant origins of highly eusocial behavior in bees (Hymenoptera: Apinae). *Syst. Biol.* 50, 194–214.

Cardinal, S., Straka, J. and Danforth, B. N. (2010). Comprehensive phylogeny of apid bees reveals the evolutionary origins and antiquity of cleptoparasitism. *Proc. Natl. Acad. Sci.* 107, 16207–16211.

Chandrasekaran, S., Ament, S. A., Eddy, J. A., Rodriguez-Zas, S. L., Schatz, B. R., Price, N. D. and Robinson, G. E. (2011). Behavior-specific changes in transcriptional modules lead to distinct and predictable neurogenomic states. *Proc. Natl. Acad. Sci.* 108, 18020-18025.

Colgan, T. J., Carolan, J. C., Bridgett, S. J., Sumner, S., Blaxter, M. L. and Brown, M. J. F. (2011). Polyphenism in social insects: Insights from a transcriptome-wide analysis of gene expression in the life stages of the key pollinator, *Bombus terrestris*. *BMC Genomics* 12, 623.

Cronin, A. L. and Hirata, M. (2003). Social polymorphism in the sweat bee *Lasioglossum (Evylaeus) baleicum* (Cockerell) (Hymenoptera, Halictidae) in Hokkaido, northern Japan. *Insectes Soc.* 50, 379–386.

Eickwort, G. C., Eickwort, J. M., Gordon, J., Eickwort, M. A. and Wcislo, W. T. (1996). Solitary behavior in a high-altitude population of the social sweat bee *Halictus rubicundus* (Hymenoptera: Halictidae). *Behav. Ecol. Sociobiol.* 38, 227-233.

Feldmeyer, B., Elsner, D. and Foitzik, S. (2014). Gene expression patterns associated with caste and reproductive status in ants: Worker-specific genes are more derived than queen-specific ones. *Mol. Ecol.* 23, 151–161.

Ferreira, P. G., Patalano, S., Chauhan, R., Ffrench-Constant, R., Gabaldón, T., Guigó, R. and Sumner, S. (2013). Transcriptome analyses of primitively eusocial wasps reveal novel insights into the evolution of sociality and the origin of alternative phenotypes. *Genome Biol.* 14, R20.

Field, J., Paxton, R. J., Soro, A. and Bridge, C. (2010). Cryptic plasticity underlies a major evolutionary transition. *Curr. Biol.* 20, 2028–2031.

Gadagkar, R. (1997). The evolution of caste polymorphism in social insects genetic release followed by diversifying evolution. *J. Genet.* 76, 167–179.

Galbraith, D. A., Kocher, S. D., Glenn, T., Albert, I., Hunt, G. J., Strassmann, J. E., Queller, D. C. and Grozinger, C. M. (2016). Testing the kinship theory of intragenomic conflict in honey bees (*Apis mellifera*). *Proc. Natl. Acad. Sci.* 113, 1020–1025.

Ghalambor, C. K., Hoke, K. L., Ruell, E. W., Fischer, E. K., Reznick, D. N. and Hughes, K. A. (2015). Non-adaptive plasticity potentiates rapid adaptive evolution of gene expression in nature. *Nature* 525, 372-375.

Grabherr, M. G., Haas, B. J., Yassour, M., Levin, J. Z., Thompson, D. A., Amit, I., Adiconis, X., Fan, L., Raychowdhury, R., Zeng, Q., et al. (2011). Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nat. Biotechnol.* 29, 644–652.

Grozinger, C. M., Fan, Y., Hoover, S. E. R. and Winston, M. L. (2007). Genome-wide analysis reveals differences in brain gene expression patterns associated with caste and reproductive status in honey bees (*Apis mellifera*). *Mol. Ecol.* 16, 4837–4848.

Harpur, B. A., Kent, C. F., Molodtsova, D., Lebon, J. M. D., Alqarni, A. S., Owayss, A. A., and Zayed, A. (2014). Population genomics of the honey bee reveals strong signatures of positive selection on worker traits. *Proc. Natl. Acad. Sci.* 111, 2614–2619.

Harrison, M. C., Hammond, R. L. and Mallon, E. B. (2015). Reproductive workers show queenlike gene expression in an intermediately eusocial insect, the buff-tailed bumble bee *Bombus terrestris*. *Mol. Ecol.* 24, 3043–3063.

Heil, M., Greiner, S., Meimberg, H., Krüger, R., Noyer, J.-L., Heubl, G., Eduard Linsenmair, K. and Boland, W. (2004). Evolutionary change from induced to constitutive expression of an indirect plant resistance. *Nature* 430, 205–208.

Hunt, B. G., Wyder, S., Elango, N., Werren, J. H., Zdobnov, E. M., Yi, S. V and Goodisman, M. A. D. (2010). Sociality is linked to rates of protein evolution in a highly social insect. *Mol. Biol. Evol.* 27, 497–500.

Hunt, B. G., Ometto, L., Wurm, Y., Shoemaker, D., Yi, S. V., Keller, L. and Goodisman, M. A. D. (2011). Relaxed selection is a precursor to the evolution of phenotypic plasticity. *Proc. Natl. Acad. Sci.* 108, 15936–15941.

Jones, B. M., Wcislo, W. T. and Robinson, G. E. (2015). Developmental transcriptome for a facultatively eusocial bee, *Megalopta genalis*. *G3* 5, 2127-2135.

Kapheim, K. M., Bernal, S. P., Smith, A. R., Nonacs, P. and Wcislo, W. T. (2011). Support for maternal manipulation of developmental nutrition in a facultatively eusocial bee, *Megalopta genalis* (Halictidae). *Behav. Ecol. Sociobiol.* 65, 1179–1190.

Kapheim, K. M., Smith, A. R., Ihle, K. E., Amdam, G. V., Nonacs, P. and Wcislo, W. T. (2012). Physiological variation as a mechanism for developmental caste-biasing in a facultatively eusocial sweat bee. *Proc. R. Soc. B Biol. Sci.* 279, 1437-1466.

Kapheim, K. M., Smith, A. R., Nonacs, P., Wcislo, W. T. and Wayne, R. K. (2013). Foundress polyphenism and the origins of eusociality in a facultatively eusocial sweat bee, *Megalopta genalis* (Halictidae). *Behav. Ecol. Sociobiol.* 67, 331–340.

Kapheim, K. M., Pan, H., Li, C., Salzberg, S. L., Puiu, D., Magoc, T., Robertson, H.M., Hudson, M.E., Venkat, A., Fischman, B.J. et al. (2015). Genomic signatures of evolutionary transitions from solitary to group living. *Science* 348, 1139–1143.

Kocher, S. D. and Paxton, R. J. (2014). Comparative methods offer powerful insights into social evolution in bees. *Apidologie* 45, 289–305.

Ledon-Rettig, C. C., Pfennig, D. W. and Nascone-Yoder, N. (2008). Ancestral variation and the potential for genetic accommodation in larval amphibians: Implications for the evolution of novel feeding strategies. *Evol. Dev.* 10, 316–325.

Leek, J. T., Scharpf, R. B., Bravo, H. C., Simcha, D., Langmead, B., Johnson, W. E., Geman, D., Baggerly, K. and Irizarry, R. A. (2010). Tackling the widespread and critical impact of batch effects in high-throughput data. *Nat. Rev. Genet.* 11, 733-739.

Leek, J. T., Johnson, W. E., Parker, H. S., Jaffe, A. E. and Storey, J. D. (2012). The SVA package for removing batch effects and other unwanted variation in high-throughput experiments. *Bioinformatics* 28, 882–883.

Leichty, A. R., Pfennig, D. W., Jones, C. D. and Pfennig, K. S. (2012). Relaxed genetic constraint is ancestral to the evolution of phenotypic plasticity. *Integr. Comp. Biol.* 52, 16–30.

Linksvayer, T. A. and Wade, M. J. (2009). Genes with social effects are expected to harbor more sequence variation within and between species. *Evolution* 63, 1685–1696.

McCarthy, D. J., Chen, Y. and Smyth, G. K. (2012). Differential expression analysis of multifactor RNA-Seq experiments with respect to biological variation. *Nucleic Acids Res.* 40, 4288–4297.

Mi, H., Muruganujan, A. and Thomas, P. D. (2013). PANTHER in 2013: modeling the evolution of gene function, and other gene attributes, in the context of phylogenetic trees. *Nucleic Acids Res.* 41, D377-86.

Moczek, A. P., Sultan, S., Foster, S., Ledon-Rettig, C., Dworkin, I., Nijhout, H. F., Abouheif, E. and Pfennig, D. W. (2011). The role of developmental plasticity in evolutionary innovation. *Proc. R. Soc. B Biol. Sci.* 278, 2705-2713.

Packer, L. (1990). Solitary and eusocial nests in a population of *Augochlorella striata* (Provancher) (Hymenoptera; Halictidae) at the northern edge of its range. *Behav. Ecol. Sociobiol.* 27, 339–344.

Pichancourt, J. B. and van Klinken, R. D. (2012). Phenotypic plasticity influences the size, shape and dynamics of the geographic distribution of an invasive plant. *PLoS One* 7, e32323.

Piperno, D. R., Holst, I., Winter, K. and McMillan, O. (2015). Teosinte before domestication: Experimental study of growth and phenotypic variability in Late Pleistocene and early Holocene environments. *Quat. Int.* 363, 65–77.

Plateaux-Quénu, C., Plateaux, L. and Packer, L. (2000). Population-typical behaviours are retained when eusocial and non-eusocial forms of *Evylaeus albipes* (F.) (Hymenoptera, Halictidae) are reared simultaneously in the laboratory. *Insectes Soc.* 47, 263-270.

Rehan, S. M. and Toth, A. L. (2015). Climbing the social ladder: the molecular evolution of sociality. *Trends Ecol. Evol.* 30, 426–433.

Santana, S. E. and Dumont, E. R. (2009). Connecting behaviour and performance: The evolution of biting behaviour and bite performance in bats. *J. Evol. Biol.* 22, 2131–2145.

Schwarz, M. P., Richards, M. H. and Danforth, B. N. (2007). Changing paradigms in insect social evolution: Insights from Halictine and Allodapine bees. *Annu. Rev. Entomol.* 52, 127–150.

Smith, A. R., Wcislo, W. T. and O'Donnell, S. (2007). Survival and productivity benefits to social nesting in the sweat bee *Megalopta genalis* (Hymenoptera: Halictidae). *Behav. Ecol. Sociobiol.* 61, 1111–1120.

Smith, A. R., Wcislo, W. T. and O'Donnell, S. (2008). Body size shapes caste expression, and cleptoparasitism reduces body size in the facultatively eusocial bees *Megalopta* (Hymenoptera: Halictidae). *J. Insect Behav.* 21, 394.

Smith, A. R., Kapheim, K. M., O'Donnell, S. and Wcislo, W. T. (2009). Social competition but not subfertility leads to a division of labour in the facultatively social sweat bee *Megalopta genalis* (Hymenoptera: Halictidae). *Anim. Behav.* 78, 1043–1050.

Smith, A. R., Kapheim, K. M., Pérez-Ortega, B., Brent, C. S. and Wcislo, W. T. (2013). Juvenile hormone levels reflect social opportunities in the facultatively eusocial sweat bee *Megalopta genalis* (Hymenoptera: Halictidae). *Horm. Behav.* 63, 1–4.

Snell-Rood, E. C., Cash, A., Han, M. V., Kijimoto, T., Andrews, J. and Moczek, A. P. (2011). Developmental decoupling of alternative phenotypes: Insights from the transcriptomes of horn-polyphenic beetles. *Evolution* 65, 231–245.

Standage, D. S., Berens, A. J., Glastad, K. M., Severin, A. J., Brendel, V. P. and Toth, A. L. (2016). Genome, transcriptome and methylome sequencing of a primitively eusocial wasp reveal a greatly reduced DNA methylation system in a social insect. *Mol. Ecol.* 25, 1769–1784.

Suzuki, Y. and Nijhout, H. F. (2006). Evolution of a polyphenism by genetic accommodation. *Science* 311, 650–652.

Van Dyken, J. D. and Wade, M. J. (2010). The genetic signature of conditional expression. *Genetics* 184, 557–570.

Waddington, C. H. (1942). Canalization of development and the inheritance of acquired characters. *Nature* 150, 563–565.

Waddington, C. (1953). Genetic assimilation of an acquired character. *Evolution* 7, 118–126.

Wang, Y., Amdam, G. V., Rueppell, O., Wallrichs, M. A., Fondrk, M. K., Kaftanoglu, O. and Page, R. E. (2009). PDK1 and HR46 gene homologs tie social behavior to ovary signals. *PLoS One* 4, e4899.

Wang, Y., Kaftanoglu, O., Siegel, A. J., Page, R. E. and Amdam, G. V. (2010). Surgically increased ovarian mass in the honey bee confirms link between reproductive physiology and worker behavior. *J. Insect Physiol.* 56, 1816–1824.

Wcislo, W. T. and Gonzalez, V. H. (2006). Social and ecological contexts of trophallaxis in facultatively social sweat bees, *Megalopta genalis* and *M. ecuadoria* (Hymenoptera, Halictidae). *Insectes Soc.* 53, 220–225.

Wcislo, W. T., Arneson, L., Roesch, K., Gonzalez, V., Smith, A. and Fernández, H. (2004). The evolution of nocturnal behaviour in sweat bees, *Megalopta genalis* and *M. ecuadoria* (Hymenoptera: Halictidae): an escape from competitors and enemies? *Biol. J. Linn. Soc.* 83, 377–387.

West-Eberhard, M. J. (2003). *Developmental plasticity and evolution*. Oxford: Oxford University Press.

Wheeler, M. M., Ament, S. A., Rodriguez-Zas, S. L. and Robinson, G. E. (2013). Brain gene expression changes elicited by peripheral vitellogenin knockdown in the honey bee. *Insect Mol. Biol.* 22, 562–573.

Whitfield, C. W., Cziko, A. M. and Robinson, G. E. (2003). Gene expression profiles in the brain predict behavior in individual honey bees. *Science* 302, 296–299.

Wilson, E. O. (1971). *The insect societies.* Cambridge: Harvard University Press.

Wilson, E. O. and Holldobler, B. (2005). Eusociality: origin and consequences. *Proc. Natl. Acad. Sci.* 102, 13367–71.

Wolda, H. and Roubik, D. W. (1986). Nocturnal bee abundance and seasonal bee activity in a Panamanian forest. *Ecology* 67, 426–433.

Woodard, S. H., Fischman, B. J., Venkat, A., Hudson, M. E., Varala, K., Cameron, S. A., Clark, A. G. and Robinson, G. E. (2011). Genes involved in convergent evolution of eusociality in bees. *Proc. Natl. Acad. Sci.* 108, 7472–7477.

Yampolsky, L. Y., Glazko, G. V. and Fry, J. D. (2012). Evolution of gene expression and expression plasticity in long-term experimental populations of *Drosophila melanogaster* maintained under constant and variable ethanol stress. *Mol. Ecol.* 21, 4287–4299.

Yanega, D. (1988). Social plasticity and early-diapausing females in a primitively social bee. *Proc. Natl. Acad. Sci.* 85, 4374–4377.

CHAPTER 5

ANCESTRAL-LIKE FORMS OF REPRODUCTIVE PLASTICITY IN LAYING WORKER HONEY BEE COLONIES

Abstract

Phenotypic plasticity plays a critical role in evolution because it contributes to the variation upon which natural selection acts. Understanding how this plasticity evolves is thus an important goal in evolutionary biology. Eusocial insects are models for studying phenotypic plasticity, as they are characterized by queen and worker castes that develop through an environmentally-determined polyphenism. Relative to the ancestral solitary state of hymenopteran eusocial lineages, queens and workers represent a decoupling of reproductive and non-reproductive behaviors, but the mechanisms by which this decoupling occurs is unknown. In some eusocial species, such as the honey bee, workers retain plasticity for reproductive phenotypes, although under typical colony conditions they remain sterile. If a honey bee colony becomes permanently queenless, however, a number of workers activate their ovaries and become laying workers (LW). Recently, it was reported that colonies of honey bees with LW contained individuals who are physiologically capable of performing both reproductive and non-reproductive behaviors, although the incidence of egg-laying in these individuals was not assessed. In this study, I utilized an automated behavioral tracking system to characterize the frequency and occurrence of egg-laying, foraging (a key non-reproductive worker behavior), and social interactions in six LW honey bee colonies. I discovered that some individuals in LW colonies do perform both reproductive (egg-laying) and non-reproductive (foraging) behaviors, much like the presumed solitary ancestor to the honey bee lineage. These "generalist" bees, appearing solitary-like individually, coexist in one nest, reminiscent of a communal social organization. However, in addition to the generalists, most individuals in LW colonies fell along a gradient of egg-laying and foraging specializations, including some highly specialized individuals that laid eggs and did not forage while others foraged and never laid eggs. This created a reproductive/non-reproductive division of labor like that observed in primitively eusocial species. Bees that engaged in this division of labor, as well as the communal-like generalist bees, participated extensively in "trophallaxis," the exchange of fluids with nutritional

and communication components. Automated behavioral tracking revealed that the trophallaxis social network in LW colonies has properties similar to those of trophallaxis networks in queenright colonies. Individual social interaction frequency and number of interaction partners varied with behavioral specialization, suggesting a possible organizing function of trophallaxis in the division of labor observed in LW colonies. Taken together, I suggest that the multiple forms of ancestral-like reproductive plasticity exhibited in LW honey bee colonies provide a glimpse into transitional stages between solitary and social living.

Introduction

Characterizing and understanding phenotypic variation is critical to understanding evolutionary processes, as this variation provides the substrate upon which natural selection acts. Social insects are models for high levels of phenotypic variation, with multiple degrees of genetic and environmentally-determined plasticity within and across species. Among hymenopteran insects, eusociality has evolved independently at least 9 times (Bourke, 2011; Cruz, 1981; Gibbs et al., 2012; Hines et al., 2007; Hughes et al., 2008; Moreau et al., 2006; Romiguier et al., 2016; Schwarz et al., 2007), and is characterized by a reproductive division of labor between queen and worker castes, representing an environmentally-determined polyphenism well-studied across lineages (e.g., Donnell, 1998; Holldobler and Wilson, 1990; Michener, 1974; Wheeler, 1986). Queens and workers of eusocial species emerge from a totipotent genome, influenced by nutrition received during development and ultimately bifurcating along axes of morphology, physiology, and behavior. Queens are specialized for reproductive functions, including mating and egg-laying, and in the complex eusocial species have ovaries with outputs orders of magnitude above their solitary ancestors. Workers, on the other hand, typically do not perform reproductive behaviors and in many cases are sterile or unable to mate, instead performing many non-reproductive behaviors in a colony. These behavioral and physiological differences arise in part through differences in gene expression and methylation patterns, which have been described in castes of many eusocial species, including honey bees (Elango et al., 2009; Feldmeyer et al., 2014; Grozinger et al., 2007; Harrison et al., 2015; He et al., 2017; Jones et al., 2017; Kucharski et al., 2008; Shi et al., 2017).

Social insects evolved from solitary ancestors, who perform both reproductive and non-reproductive behaviors in a cycle thought to be linked to the development of mature oocytes,

with cell building, provisioning, and oviposition occurring in cycles for each egg laid (Michener, 1974). A leading hypothesis relating to the evolution of reproductive castes, the Ovarian Ground Plan Hypothesis (OGPH; West-Eberhard, 1987), posits that ancestral gene networks underlying behavioral variation in this solitary reproductive cycle were decoupled into queen and worker castes, such that each caste ultimately lacks the set of traits expressed in the other. In other words, queens express only the reproductive-related behaviors expressed during the reproductive phase of the solitary cycle, while workers express only the non-reproductive behaviors. This putative decoupling would likely have involved changes in the timing and expression of existing gene networks (Linksvayer and Wade, 2005), rather than novel genes or proteins. Further specialization of castes could follow, particularly in lineages with obligate eusociality through disruptive selection and canalization of caste-related phenotypes. Indeed, genes with caste-biased expression or expression unique to one caste have been found to exhibit faster rates of protein evolution (Feldmeyer et al., 2014; Hunt et al., 2010), which may be a result of relaxed selection following eusocial evolution or a precursor to the evolution of castes (Hunt et al., 2011; Leichty et al., 2012).

The first support for the OGPH came from the observation that many species of social Hymenoptera exhibit the type of phenotypic plasticity predicted by this hypothesis. In species of primitively eusocial *Polistes* paper wasps, for example, some individuals remain on the nest as reproductives while others engage in non-reproductive tasks such as foraging (Turillazzi and West-Eberhard, 1996). More recent support for the OGPH and related hypotheses has been found in multiple lineages of social insects (Corona et al., 2013; Kapheim and Johnson, 2017; Libbrecht et al., 2018; Pamminger and Hughes, 2017; Toth et al., 2007), including highly derived species such as the honey bee (Amdam et al., 2006; Amdam et al., 2008; Graham et al., 2011; Ihle et al., 2010). Much of this support comes from correlational studies of gene expression, or manipulations of worker behaviors via social conditions or reproductive signaling pathways. These studies suggest that ancestral gene networks have been co-opted during the evolution of caste-related behaviors, in line with predictions of the OGPH. However, the mechanistic basis of this co-option is still unknown.

The honey bee has been an excellent model for understanding environmentally-mediated social plasticity, including plasticity in worker reproduction. Honey bee workers possess functional ovaries, and although they cannot mate, they can produce viable haploid eggs. In a

queenright colony, a small number (<1-2%; Ratnieks, 1993; Visscher, 1996) of workers have activated ovaries and have been found to produce up to 7% of the male eggs laid in the colony (Visscher, 1996). However, very few (~0.12%, Visscher, 1989) adult males (drones) are the product of these workers. This mismatch is due to extensive worker policing, in which workers detect and cannibalize worker-laid eggs (Ratnieks, 1993; Visscher, 1996). While rare under most queenright conditions, laying workers (LW) are frequent in situations of permanent queenlessness, when colonies lose their queen and then fail to rear a replacement queen. In these permanently queenless scenarios, up to 50% of workers may activate their ovaries (Sakagami, 1954) and some of these workers lay eggs, producing thousands of drones prior to colony death (Page and Erickson, 1988). While not all workers in these colonies lay eggs, the colony state is collectively referred to as a LW colony, and is distinct from other forms of queenless colonies due to the presence of LW.

LW colonies have long been thought to represent a collapse of colony cooperation (Malka et al., 2008), and beekeepers questioned whether LW could "function in a normal way in a colony" (Morse, 1990), displaying a general "lack of order" in the hive (Dadant & Sons, 1975). Since workers cannot mate, LW colonies also represent a terminal stage of colony life, with a dwindling population as workers die and are unable to replace themselves. However, a few lines of evidence suggest that LW behavior may have fitness benefits, including the preference for LW to oviposit in drone-sized cells (Page and Erickson, 1988) and the presence of viable sperm in LW-derived drones (Gençer and Kahya, 2011). In addition, more recent work discovered that many workers with developed ovaries in LW colonies engaged in cooperative behaviors necessary for colony function, including colony defense and foraging (Naeger et al., 2013). These results suggest that not only do LW perform multiple "normal" worker functions, but individuals in these colonies may have an expanded behavioral repertoire relative to queenright workers, similar to those of solitary or communal bees (Naeger et al., 2013). However, Naeger et al. (2013) relied on manual observations of just a subset of colony members, and primarily used ovary activation as a proxy for reproductive behavior because few bees were directly observed engaging in egg-laying behavior. It was thus not possible to know how the social organization of LW colonies was influenced by these apparently more totipotent workers.

In this study, I utilized an automated behavioral tracking system (based on Gernat et al. 2018) that employs convolutional neural networks and positional tracking, resulting in one-

second-resolution individual behavioral data on six colonies of LW honey bees. This level of resolution was necessary to obtain detailed behavioral records of all the bees in a colony, including potentially rare behavioral profiles of individual bees. Manual observation studies are limited to either scan sampling methods, which capture only a small percentage of behaviors randomized across bees, or focal sampling, in which only a few bees are able to be tracked. With the automatic tracking system, the frequency and occurrence of multiple behaviors was determined for each bee in each colony. These behaviors included egg-laying (a reproductive behavior), foraging (a commonly studied non-reproductive behavior), and trophallaxis (a social behavior between a pair of individuals, described below). Foraging has been well-studied in honey bees, and recent automatic tracking of foraging in queenright honey bee colonies provided a baseline for comparison with the results presented here (Tenczar et al., 2014). Trophallaxis, in which two bees orally transfer liquid food, proteins and small molecules (Free, 1956; Leboeuf et al., 2012; Nixon and Ribbands, 1952; Winston, 1987), was recently automatically tracked and utilized to understand communication dynamics in honey bee social networks of queenright colonies (Gernat et al., 2018). While trophallaxis is most obviously a means of food transmission in social insect colonies, it has long been hypothesized to also be a critical component of communication (Korst and Velthuis, 1982; Nixon and Ribbands, 1952). Recent work demonstrating the presence of a number of non-food molecules in trophallactic fluid provides support for this idea (Leboeuf et al., 2012; Leboeuf et al., 2016).

I used the behavioral data obtained with the automated tracking system to explore the extent of behavioral plasticity in individual members of LW colonies. I also examined the overall social organization of LW colonies with respect to division of labor and trophallaxis among colony members. Specifically, I tested the following two hypotheses: 1) laying workers act as generalists, performing both reproductive and non-reproductive behaviors, and 2) laying worker colonies have a division of labor. If hypothesis 1 is supported, it would demonstrate a surprisingly high degree of individual plasticity for a worker bee. This plasticity would normally be modulated by the social environment, where queen and brood pheromones act to inhibit ovary activation (Hoover et al., 2003; Mohammedi et al., 1998) and lead to expression of almost exclusively non-reproductive behaviors among workers. This degree of flexibility in worker behavior despite a long evolutionary history of eusociality and suppression of reproduction would thus suggest that relatively simple switches in gene regulation underlie the decoupling of

behaviors associated with the evolution of caste.

If, however, a division of labor exists among colony members who are all physiologically capable of both reproductive and non-reproductive behaviors (hypothesis 2), LW colonies may resemble a primitively eusocial state, a potential intermediary stage in the evolution of complex eusociality from solitary ancestors (Rehan and Toth, 2015). While transitions from solitary to eusocial behavior may not have included stable intermediate stages, many incremental steps are involved in eusocial origins, suggesting that a "social ladder" of transitional stages is possible, at least in some of the independent evolutions of eusociality (Evans and West-Eberhard, 1970; Rehan and Toth, 2015). Behavioral decoupling among LW linked to variation in reproductive activity would provide additional support for an extension of the OGPH in honey bees, where subcaste specialization is suggested to arise through co-option of reproductive gene regulatory networks (Amdam et al., 2006; Page and Amdam, 2007; Page et al., 2006).

Elucidating the individual behavioral plasticity and colony-level organization of colonies with egg-laying worker honey bees provides the potential for important new insights on how this highly derived species can be used to better understand the evolution of eusociality. My results provide insights into the utility of LW honey bee colonies for studies of transitions in social organization, and open doors to future approaches to further understand the evolution and molecular architecture of castes.

Methods

Bees and colony setup

Source colonies

Honey bee colonies were maintained according to standard beekeeping practices at the University of Illinois Bee Research Facility in Urbana, Illinois. One-day-old adult worker bees were obtained by removing sealed frames of late-stage pupae from source colonies and housing them in an incubator inside emergence cages at 34 °C and 50% relative humidity. Bees were removed from frames daily to collect adults less than 24 hours old.

Prior to establishing the colonies of barcoded bees, multiple colonies were screened for worker egg-laying potential by creating Plexiglas cages with 50-100 one-day-old workers. Cages contained small pieces of 3D-printed honeycomb (similar to Fine et al., 2018) to provide a standardized location for workers to lay eggs, as well as 50% sucrose solution and pollen paste

(45:45:10 ratio by weight of pollen, honey, and water) provided *ad libitum* and refreshed daily. Cages were monitored daily to count eggs. We found, as in other studies, variation in the timing and extent of laying worker development among different source colonies (Appendix Figure A.5), reflecting genotypic and/or environmental differences in laying worker potential (Miller and Ratnieks, 2001; Page and Robinson, 1994; Robinson et al., 1990; Velthuis, 1970). Source colonies for experiments were chosen, when possible, from among those screened that displayed high levels of worker egg-laying in cages within 14 days.

Experimental colonies A-C were established from naturally-mated, *Apis mellifera ligustica* colony sources. Experimental colonies D-F were established from a mix of two source colonies each headed by an artificially inseminated queen, who had been single-drone inseminated (SDI) and was of either *Apis mellifera ligustica* or *Apis mellifera carnica* origin (queen rearing and inseminations performed by Sue Cobey- Honey Bee Insemination Service; Washington State University). One-day-old worker bees were obtained from these colonies as above, with late-stage pupae incubated and 0-24-hour-old workers collected each day prior to barcoding. A total of 800 bees were used for each experimental colony, collected and barcoded over 1-2 days from the emerging incubator frames (Appendix Table B.1).

Barcoding bees

Bees were barcoded with custom "bCodes" as in Gernat et al. (2018). Unique sets of bCodes were used to differentiate bees barcoded on different days, as well as to differentiate bees from different source colonies in colonies D-F. Briefly, workers were anesthetized on ice and then positioned using soft forceps (BioQuip Products). A small drop of Loctite Super Glue Gel Control (Henkel) was applied to the center of the thorax of each bee, followed by a bCode positioned with its orientation vector parallel to the anteroposterior axis of the bee. Bees were carefully placed in plastic dishes until they recovered from cold anesthetization, at which point the glue was dry. After waking, all bees were placed in a large container with Fluon®-coated walls (Insect-a-Slip, BioQuip) where honey was provided *ab libitum* until placement into the hive. At the end of each barcoding day, bees were carefully transferred into a custom observation hive, described below.

Behavioral tracking

Hive monitoring

Barcoded bees were housed in a glass-walled observation hive with a one-sided

plastic honeycomb frame, as in Gernat et al. (2018). Bees were unable to access the back side of the hive, and could exit the hive through a plastic tube to the outside. Colonies were maintained in a dark room with a heater and humidifier that kept the room at approximately 32°C and 50% relative humidity.

Infrared lights (not visible to bees) were used to illuminate the hive from both the front and back during the capturing of hive images, controlled by a CTL-IO-4 I/O Module (Smart Vision Lights). Images were acquired at one-second resolution with a monochrome Prosilica GX6600 machine vision camera (Allied Vision) fitted with a Nikkor AF 135 mm f/2 D DC prime lens (Nikon). Additional detail about image acquisition can be found in Gernat et al. (2018). Images were saved to a redundant array of independent disks, then copied onto a computing cluster (Biocluster, UIUC) for analysis after the end of each experimental recording period.

Entrance monitoring

Colonies of barcoded bees were given access to the outside via a tube connected through an exterior wall of the Bee Research Facility (UIUC) to an entrance equipped with an automated flight activity monitor. This monitor included a maze to slow down incoming and outgoing bees, and a Raspberry Pi camera (5 megapixel v1.3, Adafruit) that imaged the maze twice per second from 07:00 until 19:00 daily. The camera was controlled by a Raspberry Pi 2B computer running the Raspian 8 operating system using a custom shell script and the raspistill program to record images.

Barcode detection

Barcodes were detected in hive images as in Gernat et al. (2018), and filtered to facilitate subsequent behavioral analyses. Filtering involved removal of potential tracking errors, including removal of barcodes that did not pass error detection or error correction. In addition, records for barcodes that were read twice in the same image were removed, as were hive image records of the same barcode identified more than 5 cm/second between successive detections, which are likely to be misidentifications. An average of 94.51% of detections remained after these filtering steps (range across colonies: 91.94-97.11%). Finally, the time of death of each bee was estimated using the last time she was observed for at least 4 minutes during a 5-minute window above the third row of honeycomb cells from the bottom of the hive, which is where dead bees tended to accumulate prior to being removed by other bees (Gernat et al., 2018). Records for bees

following their time of death were filtered out as to not influence behavioral scores by normalizing over times in which the bee was not alive.

In entrance monitor images, barcodes were similarly detected as in hive images, but with parameters adjusted to adapt to the images produced by the entrance monitor. In addition, fast-moving bees were not filtered out, because bees do move quickly through the entrance monitor and due to the relatively small number of bees that fit into the maze, spurious fast movement due to bCode decoding errors is unlikely.

Egg-laying detector

Annotated image library

Hive images from three experimental colonies and across 12 different days were used for manual annotation of egg-laying events. The software Fiji (Schindelin et al., 2012) was used to mark the bCode positions of all workers laying eggs in an initial set of 1500 hive images, followed by an additional set of 782 images, each annotated by three independent observers. After the initial identification of egg-laying bees in these images, the two seconds before and after each egg-laying event were also annotated for those bees. Bees not marked as laying eggs with visible bCodes were considered non-egg laying for training of the CNN, below.

CNN training and performance estimation

Two convolutional neural networks (CNNs) were trained on the annotated egg-laying images, using TensorFlow™ (Abadi et al., 2016). The first CNN used images cropped to include just a small rectangular region behind the barcode of each bee. For egg-laying bees, these images show the honeycomb, because their abdomen is backed into the comb and thus not visible. For non-layers, these images show the abdomen. The CNN was trained to differentiate between these two cases. The second CNN was applied to images of bees that were identified as potential egg-layers by the first CNN. It used slightly larger images that showed the entire bee and was trained to use information about the bee's posture and her immediate surroundings to identify false positives, which were subsequently filtered out.

Application of a CNN to an image results in a score between 0 and 1 that reflects the likelihood of that image showing the event of interest. Deciding whether a score is sufficiently high for assuming that the event took place involves thresholding that score. To choose thresholds for each CNN score and a minimum egg-laying duration, a calibration set of images, which were not used for training the CNNs, was used to estimate the performance of the egg-

laying detector for different threshold combinations. Thresholds were chosen from this calibration set to minimize false positives, then were applied to an independent test set of images that had also never been seen by the detector to obtain unbiased performance values. In addition, a bee was required to pass these thresholds for at least 3 seconds, as a shorter time is unlikely to be representative of a real egg-laying event (Page and Erickson, 1988; Sakagami, 1958; Velthuis, 1970, Jones, pers. obs.). Based on performance estimation on the test set of images, the egg-laying detector had the following performance metrics: 99.71% accuracy, 100% positive predictive value, 99.71% negative predictive value, and 35.39% sensitivity. Minimizing false positives came at a cost to sensitivity, but bees who lay eggs will likely do so more than once over the course of tracking (honey bees possess multiple ovarioles, each of which can develop eggs simultaneously). Egg-laying detections were further aggregated into events. Subsequent detections that occurred within 10 seconds or 11.2 mm (the width of two honeycomb cells) of one another were assumed to belong to the same egg-laying event and were merged.

Trophallaxis detector

Annotated image library

Image library *L1* was used to obtain images for trophallaxis detector training (Gernat et al., 2018). Briefly, *L1* consists of approximately 1,000 images, each showing a pair of bees engaged in trophallaxis, and roughly 40,000 images, each showing a pair of bees in the proper position for trophallaxis, but not performing the behavior. For training purposes, only 1,000 of the latter class of images were used such that equal numbers of trophallactic and non-trophallactic pairs were included. Additional images were obtained from an extended version of image library *L2* (Gernat et al., 2018). This library consists of images showing all pairs of bees that are in the proper position for trophallaxis in 100 randomly chosen triples of successive hive images (i.e., three chronological images of the same bee pair). The extension consisted of an additional 100 annotated image triples as well as 100 and 200 randomly chosen triples obtained from the hive image data sets of two additional colonies.

CNN training and performance estimation

TensorFlow™ (Abadi et al., 2016) was used to train two CNNs for detecting the occurrence of trophallaxis and for determining the trophallactic role (donor or recipient) of each trophallaxis partner. Both CNNs used as input a small image showing only the head and mouthparts of potential trophallaxis partners that had been identified with a permissive version

of the geometric filter described in Gernat et al. (2018). The first CNN was trained to distinguish images showing trophallaxis from images not showing that behavior. The second CNN was trained to identify donor and recipient bees in images showing trophallaxis.

To choose thresholds for each CNN score and parameters for the geometric filter, a calibration set of images, which was not used for training the CNNs, was used to estimate the performance of the detector for different threshold and parameter combinations on image triples (requiring that trophallaxis be detected in all three images), and the values maximizing the product of detector sensitivity and positive predictive value were chosen. Lastly, the detector was parameterized with these values and applied to an independent test set of image triples that it had never seen to obtain unbiased performance values. Parameters used for the detector resulted in the following performance metrics: 88.7% sensitivity, 99.6% specificity, 90.4% positive predictive value, 99.6% negative predictive value, and 88.9% accuracy in determining trophallactic role (donor or receiver) of each bee.

Filtering and annotation of entrance data

Raw detections of bees in the entrance were filtered such that a bee must traverse at least one-third the distance of the entrance monitor to be counted. Traversals that occurred within 10 seconds of each other were merged into a single event. These traversal events were then determined to be incoming or outgoing based on the positional coordinates of the bee at the start and end times of each event. Flight activity was inferred from series of outgoing and incoming events.

Incoming trips were additionally annotated with trophallaxis data to determine whether an incoming forager likely returned with nectar or water. If a bee was a trophallaxis donor within 5 minutes after returning from a trip, with no trophallaxis reception prior to the donation, she was annotated as having a donation event associated with that incoming event. To be categorized as a forager on a given day, a bee had to make at least two trips that were either 5 minutes in length (to remove orientation flights, which are typically shorter in length, after Ribbands, 1952; Robinson, 1985; Sekiguchi and Sakagami, 1966; Vollbehr, 1975; Winston and Katz, 1982) or were followed by a trophallaxis donation. For colonies D-F, all incoming trips were additionally manually annotated for pollen on the hindlegs of returning bees. Incoming trips where bees were visibly seen with pollen were almost exclusively 5 minutes or longer in length

(98.8% of trips annotated with pollen were longer than 5 minutes; Appendix Figure A.6), suggesting that the 5 minute threshold did not often eliminate true pollen foraging trips.

Specialist and generalist scores

In order to characterize the activity of egg-laying and foraging for each bee, two behavioral scores were created. The specialist score describes how specialized an individual was on either egg-laying (scores near -1) or foraging (scores near +1) relative to other bees in the colony; bees that consistently performed both egg-laying and foraging, or that performed neither behavior, have specialist scores near 0. The generalist score ranges from 0 to 1 and describes the degree to which an individual performed both egg-laying and foraging behaviors, differentiating bees with specialist scores near 0 based on the performance (or not) of egg-laying and foraging. Scores were created by first counting the number of egg-laying and foraging events per day. Bees were then ranked for each behavior relative to other bees in the colony on the same day, with tying ranks being assigned the minimal rank (e.g., if three bees were tied between the 4^{th} and 8^{th} ranked bees, they all received a rank of 5). Ranks were then normalized by dividing by the maximum rank, so that all ranks were in the range [0,1]. The normalized rank space for each bee (i.e. normalized egg-laying rank and normalized foraging rank) was then mapped to behavioral scores (and corresponding color space) using the following formulae in polar coordinates (ρ,Θ) on the two-dimensional rank space: generalist score = $(1/2)\rho^2\sin^4 2\Theta$, specialist score= $\sin(\Theta-\pi/4)\rho^4\cos^4 2\Theta$. Note that the numerical value of the scores has no biological meaning, but are simply a mapping from rank space to the space of colors as shown in Appendix Figure A.7.

Ovary dissections

For a subset of bees with either extreme specialist scores (top specialized foragers and egg-layers) or high generalist scores, abdomens were dissected to assess ovary activation. Abdomens of each bee were carefully removed on dry ice and incubated for 16 hours at -20°C in RNA-later ICE (Life Technologies). Ovaries were imaged and assessed for ovary development using a 1-5 scale adapted from Hess (1942), where a score of 3-5 indicates ovary activation. Ovary scores, as well as number of ovarioles as determined from dissections, are given in Appendix Table B.2.

Results

Colony-level temporal patterns of activity

All colonies collectively exhibited reproductive (egg-laying), non-reproductive (foraging), and trophallaxis behaviors during the tracking period. Average rates of egg-laying, foraging, and trophallaxis for each colony are presented in Table 5.1.

Egg-laying was observed in all colonies by the time bees were 15 days old, suggesting rapid development of ovaries by some individuals (Fig. 5.1A). With the exception of colony D, there was no obvious rhythmicity in egg-laying behavior, and egg-laying occurred throughout the day and night. Spikes of entrance activity were observed in the early days of colony development (Fig. 5.1B), which may be associated with cleansing flights or orientation flights of foraging-destined bees (Winston, 1987). Trophallaxis was observed in all colonies at rates similar to those seen in queenright colonies (Table 5.1; Gernat et al., 2018; Walton and Toth, 2016).

Classification of bees into behavioral groups

Each bee was classified daily into one of four groups: layers, foragers, generalists, and others. Layers were defined by a minimum of 2 egg-laying events per day, and foragers were defined by performance at least 2 foraging trips per day with evidence of foraging success (see Methods; typical foragers take multiple trips per day, Tenczar et al., 2014). Generalists met both layer and forager criteria on a given day, while "other" bees did not meet the minimum criteria for either behavior.

Colonies exhibited large variation in the proportion of bees within each group, particularly across time in colonies A and D for which earlier timepoints were recorded (Fig. 5.2). With the exception of colony F, a higher proportion of workers fell into the layer category than the forager category. However, the distribution of eggs laid was highly skewed (Fig. 5.3, Appendix Fig. A.8): on average, 8.4±0.6% of bees performed 50% of all egg-laying events (Table 5.2). This skew in egg-laying is greater than the skew in queenright foraging effort previously published (~20%; Tenczar et al., 2014), as well as foraging effort in the queenless colonies reported here (Fig. 5.3 and Table 5.2; on average, 12.1±0.5% of bees performed 50% of foraging trips across colonies and days). Generalists were somewhat rare, with an average of 10.8% of bees performing both egg-laying and foraging on the same day at least once during the final 7 days of tracking (range: 4.5-17% per colony). Most of these

generalists did not consistently perform both behaviors daily; across the six colonies, only 45 bees (1.3%) were classified as generalists for 3 or more days. However, removing the requirement that egg-laying and foraging occur on the same day, 600 bees across all six colonies (16.8%) performed both egg-laying and foraging behaviors between days 15 and 21. In comparison, 54% of all bees were observed laying eggs on at least one day, and 28% of all bees were observed foraging on at least one day.

Influence of worker source colony on behavior

Bees in colonies D-F were obtained from two different source colonies, each of which was headed by a queen inseminated with the semen of a single drone (SDI). As such, workers within each source colony are highly genetically related compared with workers from a naturally mated queen colony (SDI workers have an average relatedness of 0.75 due to haplodiploidy). Using these SDI colonies provided a stronger genetic signal to better explore whether the genetic and environmental differences between colonies would lead to segregation of behaviors when mixed into the same queenless environment. In colonies D and E, which were replicates of the offspring of the same two SDI queens, the classification of bees differed by genotype; one source genotype (SDI 2) contained the majority of layers, while the other (SDI 1) contained the majority of foragers (Fig. 5.2). However, layers and foragers were identified in both source colony backgrounds, suggesting multiple sources of variation contribute to the behavioral phenotypes of individual bees. In colony F, the two SDI source colony progeny contributed more equally to egg-laying and foraging phenotypes (Fig. 5.2, SDIs 3 and 4).

Individual-level consistency in behavior

Many bees specialized on either foraging or egg-laying throughout the week-long period all colonies were tracked (Fig. 5.4). Ovary dissection of specialized egg-layers and foragers revealed that while 100% of the specialized egg-layers had active ovaries, only 54% of the specialized foragers had ovaries in an activated state at the end of the tracking period (Appendix Table B.2). While some bees performed both egg-laying and foraging throughout the experiment, rarely were these bees persistent in their behavioral generalization. Extreme generalists, bees who were classified as generalists for at least 3 of the final 7 days (n=45, Fig. 5.5), had a median lifetime generalist score of 0.286, compared with the population average of 0.0110 (n=3530; t-test p=2.56e-06). Ovary dissection of a subset of these generalists (N=5) confirmed the presence of active ovaries.

As before, colonies D and E showed an influence of source colony on behavior, with the majority of specialized layers coming from SDI 2 (white bars alongside heatmaps in Fig. 5.4), while most specialized foragers were from SDI 1 (black bars). While in colonies D and E source colony was predictive of behavioral specialization, colony F shows a more equal mix of both source SDIs contributing to each behavioral group.

Engagement of reproductive workers in the trophallaxis social network

Overall, trophallaxis rates in queenless LW colonies (Table 5.1) were similar to those reported in queenright colonies (~1-2 per bee per hour; Gernat et al. 2018) and laboratory assays of queenright workers (~0.4-1.2 per bee per hour; Walton and Toth 2016). Egg laying activity was positively correlated with trophallaxis (Fig. 5.6a); bees that laid more eggs on a given day tended to engage in more trophallaxis events (Spearman's rank correlation test (SRCT); $\rho=0.48$, $p<0.0001$), and with a larger number of partners (SRCT; $\rho=0.45$, $p<0.0001$). In contrast, foraging activity was not similarly correlated with trophallaxis (Fig. 5.6B); daily per-bee foraging levels had no significant correlation to the overall number of trophallaxis interactions per day (SRCT; $p=0.2463$), and the number of interaction partners was only very weakly associated with foraging activity (SRCT; $\rho=0.02$, $p=0.018$).

The lifetime specialist score of each bee was also correlated with the number of lifetime trophallaxis interactions and number of interaction partners per bee (Fig. 5.7A). More specialized egg layers (bees with negative specialist scores) tended to engage in more trophallaxis interactions (SRCT; $\rho= -0.37$, $p<0.0001$) and with more partners (SRCT; $\rho= -0.29$, $p<0.0001$). Individual lifetime generalist scores were also positively correlated with the number of lifetime trophallaxis interactions (SRCT; $\rho= 0.33$, $p<0.0001$) and number of interaction partners (SRCT; $\rho= 0.32$, $p<0.0001$) (Fig. 5.7B), suggesting that bees consistently engaged in both egg-laying and foraging, like the egg-laying specialists, are more active in the social network.

Specialized layers engaged in more trophallaxis interactions both as receiver (SRCT; $\rho= -0.40$, $p<0.0001$; Fig. 5.8A) and as donor (SRCT; $\rho= -0.24$, $p<0.0001$; Fig. 5.8B). In addition, specialist score was weakly correlated with reception:donation ratio, with specialized layers showing slight biases for reception during trophallaxis interactions (Fig. 5.8C).

Discussion

In this chapter, I demonstrate that individuals within queenless laying worker (LW)

honey bee colonies display a unique colony organization characterized by the presence of multiple ancestral-like forms of reproductive plasticity. This includes some individuals that share a nest while expressing extensive behavioral plasticity by performing both reproductive and non-reproductive tasks, much like the presumed solitary ancestors of eusocial bees (Michener, 1974; Wilson, 1971). Other individuals participate in a division of labor for reproductive and non-reproductive behaviors, similar to what is observed in primitively eusocial species of bees and wasps (Donnell, 1998; Turillazzi and West-Eberhard, 1996). The Ovarian Ground Plan Hypothesis (OGPH) posits that the evolution of eusocial castes from a generalist solitary lifestyle involved the decoupling of reproductive and non-reproductive behaviors. LW colonies may thus represent what one would see in the earliest stages of eusocial evolution, where solitary-like individuals co-occur with individuals that begin to decouple these traits.

Based on automatic behavioral tracking of LW colonies, a subset of workers (generalists) performed both reproductive (egg-laying) and non-reproductive (foraging) behaviors. This finding is consistent with previous work which discovered that many queenless workers performing non-reproductive behaviors such as foraging and colony defense had developed ovaries (Naeger et al. 2013). While this previous study utilized manual observations and primarily relied on ovary development as a proxy for reproductive behavior, I used a high-throughput automatic behavioral tracking system to study egg-laying behavior directly. My results confirm that reproductively active workers often also engage in non-reproductive behaviors, including foraging. While my study was limited in the number of behaviors tracked, results from Naeger et al. (2013) support the idea that several other non-reproductive behaviors, including colony defense, wax building, and brood care, may be performed by some egg-laying workers.

Workers in LW colonies displayed continuous variation in behavior, with both extreme specialists for a single behavior as well as generalists who performed multiple behaviors within short time windows. In the case of both foraging and egg-laying, I identified high levels of skew with respect to performance of these tasks. Like a previous study that utilized RFID technology to track large numbers of foraging bees (Tenczar et al., 2014), I identified a small number of "elite" foragers who performed a disproportionate amount of foraging effort relative to the rest of the foraging workforce. This skew among foraging effort in LW colonies was even greater than that reported in queenright colonies in Tenczar et al. (2014), reflecting an expanded range of

foraging effort variation among foragers in LW colonies. Even more extreme was the skew in egg-laying effort, with a large majority of egg-laying events performed by a small number of elite layers, suggesting there may be strong social inhibition of egg-laying, similar to the proposed social inhibition of foraging (Beshers et al., 2001; Leoncini et al., 2004). Social inhibition is pervasive in insect colonies, influencing the development of castes, behavioral maturation, and the performance of tasks among individuals (Wilson, 1971). Collectively, these results point to a high degree of division of labor present in LW colonies, with individual variation in both foraging and egg-laying efforts across bees. This is consistent with variation in activity levels observed across social insect species (Beverly et al., 2009; O'Donnell and Jeanne, 1990; Oster and Wilson, 1979), suggesting that LW colonies are governed by similar organizing principles despite their somewhat atypical origins.

My finding of a division of labor in LW colonies, with behavioral specialization similar to that seen in queenright colonies, supports the idea that LW honey bees display a form of colony organization that is potentially adaptive (Page and Erickson, 1988), as opposed to one of chaos and competition, which has long been thought to characterize LW colonies (Morse, 1990; Ratnieks and Wenseleers, 2008; Ratnieks et al., 2006; Dadant & Sons, 1975; Wenseleers and Ratnieks, 2006). A chaotic colony state may be expected to contain a large number of generalists, acting individually to perform multiple behaviors within the colony. Instead, I found that while generalists do exist in LW colonies, they are rare relative to specialized bees engaged in a division of labor between reproductive and non-reproductive tasks. Division of labor is widely assumed to contribute to the tremendous success of social insects (Chittka and Muller, 2009; Holldobler and Wilson, 1990; Wilson, 1985), and behavioral specialization has been shown to increase efficiency and fitness in some social groups (Beshers and Fewell, 2001; Pruitt and Riechert, 2011; Trumbo and Robinson, 1997). Although worker honey bees cannot mate and therefore cannot produce fertilized (female) offspring, the haplodiploid sex determination system of honey bees and other hymenopteran insects results in the production of viable male offspring from unfertilized eggs. In LW colonies, specialization along a reproductive/non-reproductive axis may contribute to the production of haploid males prior to the death of workers, with specialized foragers collecting food for these developing drones while specialized egg layers work to produce thousands of drones synchronously in these terminal colonies (Page and Erickson, 1988). These LW-produced drones have viable sperm (Gençer and Kahya, 2011), and

therefore may lead to one final shot at fitness in the terminal stages of a permanently queenless honey bee colony. The production of drones by workers in LW colonies is similar to that observed in bumblebees, where worker competition over male production is a normal part of the colony cycle after queen death (Cnaani et al., 2002; Free, 1955), or even prior to queen death in some species (Velthuis and Duchateau, 2011).

LW colonies also exhibited signatures of a social interaction network similar to those reported in queenright honey bee colonies, with a large number of bees participating in trophallaxis. In a trophallaxis interaction, liquid food is transferred between a pair of individuals (Free, 1956; Wilson, 1971), as well as a number of small molecules which may contribute to communication within a social colony (Leboeuf et al., 2012; Leboeuf et al., 2016; Nixon and Ribbands, 1952). These interactions therefore have the potential to bias resources toward some individuals more than others, with dominant individuals receiving more trophallaxis than other individuals in the colony. Reproductively dominant individuals have been shown to receive more trophallaxis in species of bees and ants, both under queenright as well as queenless conditions (Bourke, 1988; Korst and Velthuis, 1982; Pardi, 1948; Sommeijer and Van Veen, 1990). Consistent with these findings, I found that bees engaged in the highest levels of personal reproduction in LW colonies were more connected socially to other bees in the colony; they exhibited greater numbers of trophallaxis interactions and with a greater number of interaction partners relative to other bees. However, the trophallaxis receiving bias was relatively weak; egg-laying individuals were both more likely to be donors as well as receivers of trophallaxis. This suggests that either layers are not socially dominant in LW colonies, or that dominance in a LW colony is manifested in ways other than through a trophallaxis receiving bias. Alternatively, reproductively active bees in a LW colony may act as donors in trophallaxis interactions as a mechanism of appeasement during aggressive interactions (Liebig et al., 1997). It is also important to note that a comparison between layers and foragers for behaviors that occur in the hive, such as trophallaxis, is potentially biased due to the time foragers spend outside of the hive. However, generalists, who both lay eggs and forage, also showed increased trophallaxis and numbers of trophallactic partners, suggesting that despite time spent outside, bees that engage in egg-laying may interact more than non-layers. This idea is further supported by the many bees with low specialist and generalist scores that also showed low numbers of interactions and few interaction partners, despite spending all of their time inside the hive (based on the absence of

flight activity).

The presence of both a reproductive division of labor as well as solitary- or communal-like generalist bees, all interacting socially within the same colony environment, suggests that LW colonies provide a glimpse into a transitional stage between solitary life and incipient eusociality. Among individuals in a LW colony, some perform both reproductive and non-reproductive behaviors, while others exhibit decoupling of these behaviors as is typical in honey bee colonies. This may be similar to a transitional stage in social evolution, where individual variation in behavior leads to the coexistence of multiple forms of social and reproductive plasticity within a group of individuals. For example, a group of individuals sharing a nest (e.g., communal bees or wasps; Abrams and Eickwort, 1981; Danforth, 1989; Kukuk and Crozier, 1990; McCorquodale, 1989; McCorquodale and Naumann, 1988) all may engage in both reproductive and non-reproductive behaviors, but they may vary in their probability of performing these different tasks due to variation in pre-imaginal nutrition, genotype, or environmental conditions. This may lead to a weak division of labor, with more socially dominant individuals exhibiting more reproduction and subordinate individuals performing more non-reproductive behaviors. According to the OGPH (West-Eberhard, 1987; West-Eberhard, 1996), this weak division of labor could be strengthened through disruptive selection until reproductive and non-reproductive behaviors are entirely decoupled, as seen in eusocial species with obligately sterile worker castes. Initially, this decoupling may be primarily behavioral, with little to no difference in physiology between individuals that do and do not reproduce. Over evolutionary time, decoupling may also involve physiological mechanisms, extending the separation between castes. Naeger et al. (2013) found that foraging bees in LW colonies were just as likely to possess activated ovaries as those bees not foraging, suggesting a release of this decoupling of foraging from reproductive physiology in LW. However, in this study, I found that while many foragers did have activated ovaries, only about half of those at the extreme of the behavioral spectrum who specialized on foraging behavior had activated ovaries, and rarely were those ovaries developed enough to produce mature ooctyes. Individuals at the behavioral extremes may therefore represent the height of decoupling within the LW phenotype, while generalists (and less specialized bees) may be more representative of ancestral levels of behavioral and physiological flexibility. Thus, collectively, a LW colony exhibits multiple forms of social plasticity, suggesting that worker honey bees possess a latent and broad range of

phenotypes that are typically repressed by queen and brood pheromones (Hoover et al., 2003; Mohammedi et al., 1998), but are released given certain environmental and social conditions.

The behavioral plasticity observed in queenless LW colonies is surprising given the long divergence time of honey bees from their solitary ancestor (~75 my, Branstetter et al., 2017), and is another example of phenotypes that are latent in specialized social insects but inducible under extreme conditions (Simola et al., 2015; Wilson, 1980). In this particular case, the results for LW honey bees support hypotheses that mechanisms underlying reproductive division of labor may involve reversible changes in gene regulation (Linksvayer and Wade, 2005; West-Eberhard, 1987; West-Eberhard, 1996) . The unique social organization and reproductive plasticity observed in LW opens the door to studies of how ancestral gene regulatory networks are modified and selected upon during the evolution of eusocial castes. Molecular studies of queenless LW colonies may not only provide insights into how caste-related traits are regulated, but also allow mapping of the molecular underpinnings of plasticity as it relates to variation in social organization across the social insects. This is the subject of Chapter 6.

Acknowledgments

I would like to thank Tim Gernat, Tobias Jagla, and Vikyath Rao for their contributions as collaborators in the research described in this chapter.

Figures and Tables

Figure 5.1

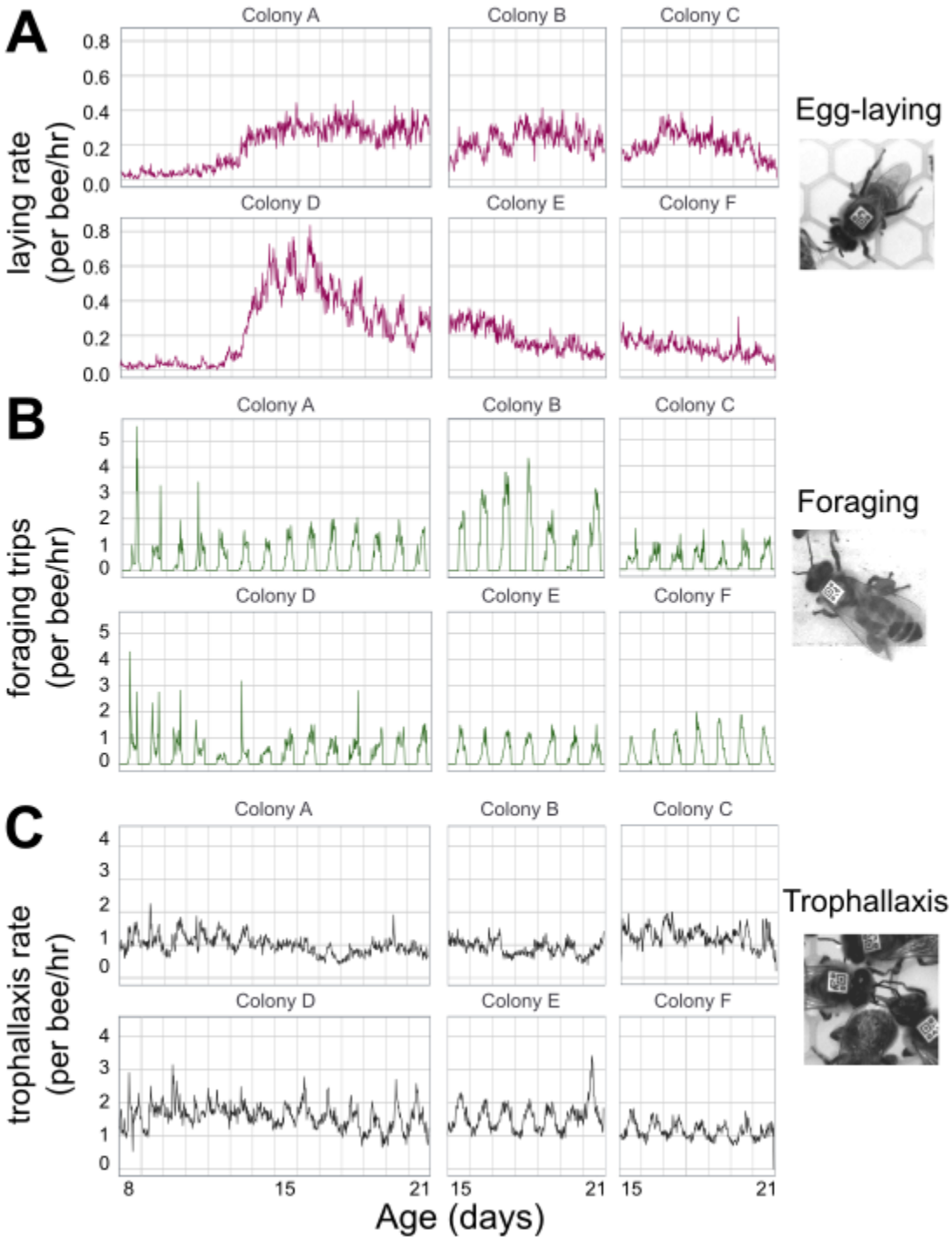


Figure 5.1. Normalized (per bee per hour) (A) egg-laying, (B) foraging, and (C) trophallaxis rates for six laying worker colonies, A-F.

Figure 5.2

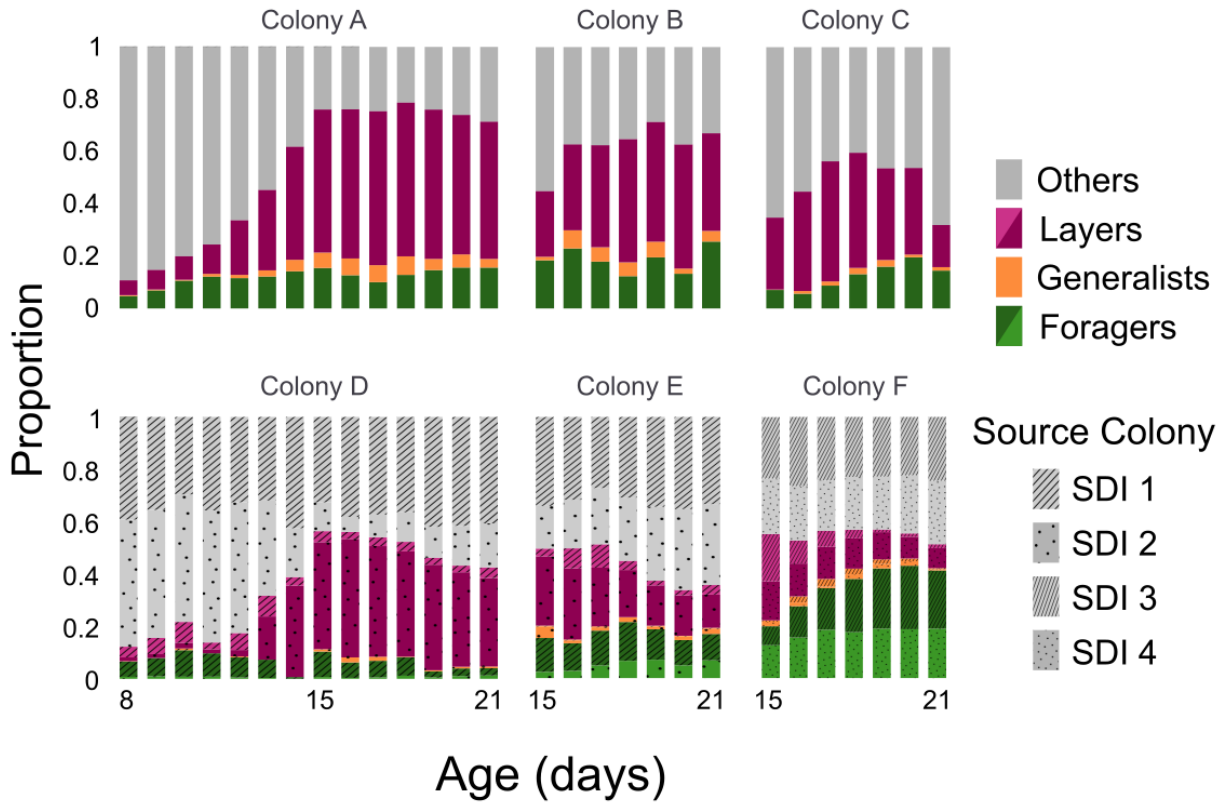


Figure 5.2. Proportion of bees alive each day categorized as layers (purple), foragers (green), generalists (orange), or others (gray). For colonies D-F, two source colonies headed by queens inseminated by semen from a single drone (single drone inseminated, SDI) were mixed. Source colonies are indicated by hashing and dots, as well as different hues for layers and foragers.

Figure 5.3

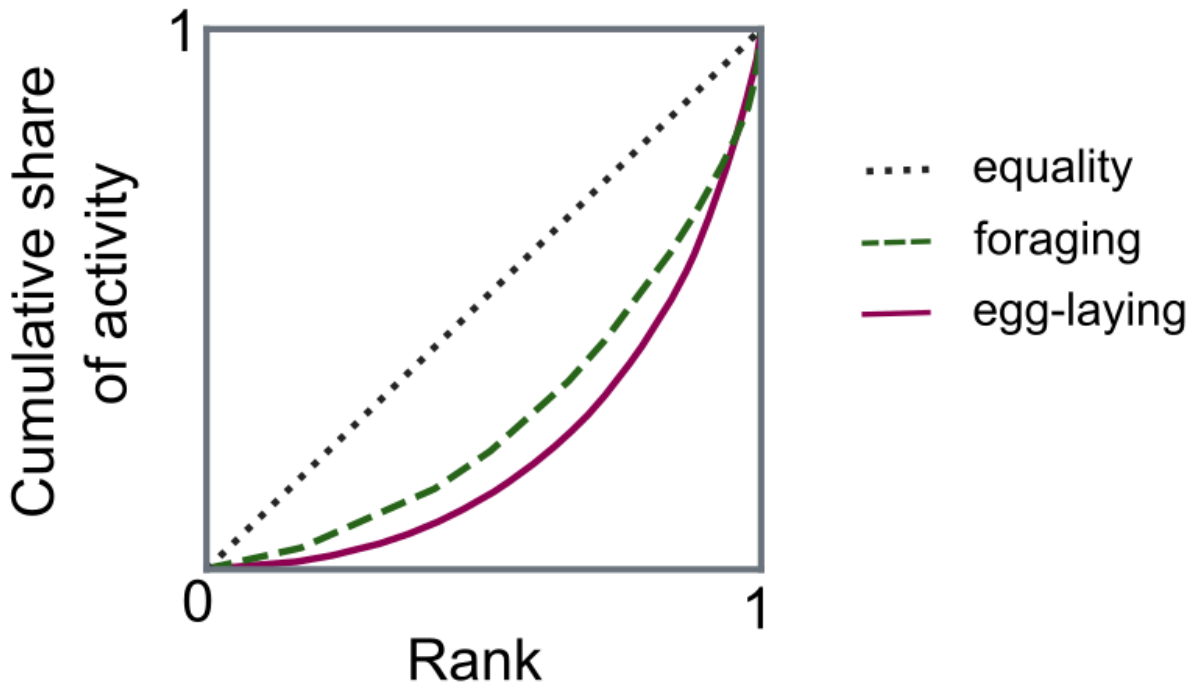


Figure 5.3. Lorenz curve for a representative colony and day. Individual bees were ranked by the number of foraging or egg-laying counts per day, and the fraction of each bee's contribution to the total activity of the colony was cumulatively plotted. If all bees in the colony contributed equally to a given behavior, the curve for that behavior would fall along the dotted "equality" line. Both foraging (green, dashed line) and egg-laying (purple, solid line) activities were unequally distributed among bees, with the level of inequality proportional to the area between the equality line and the behavioral curve. Lorenz curves for all days in all colonies are shown in Appendix Figure A.8.

Figure 5.4

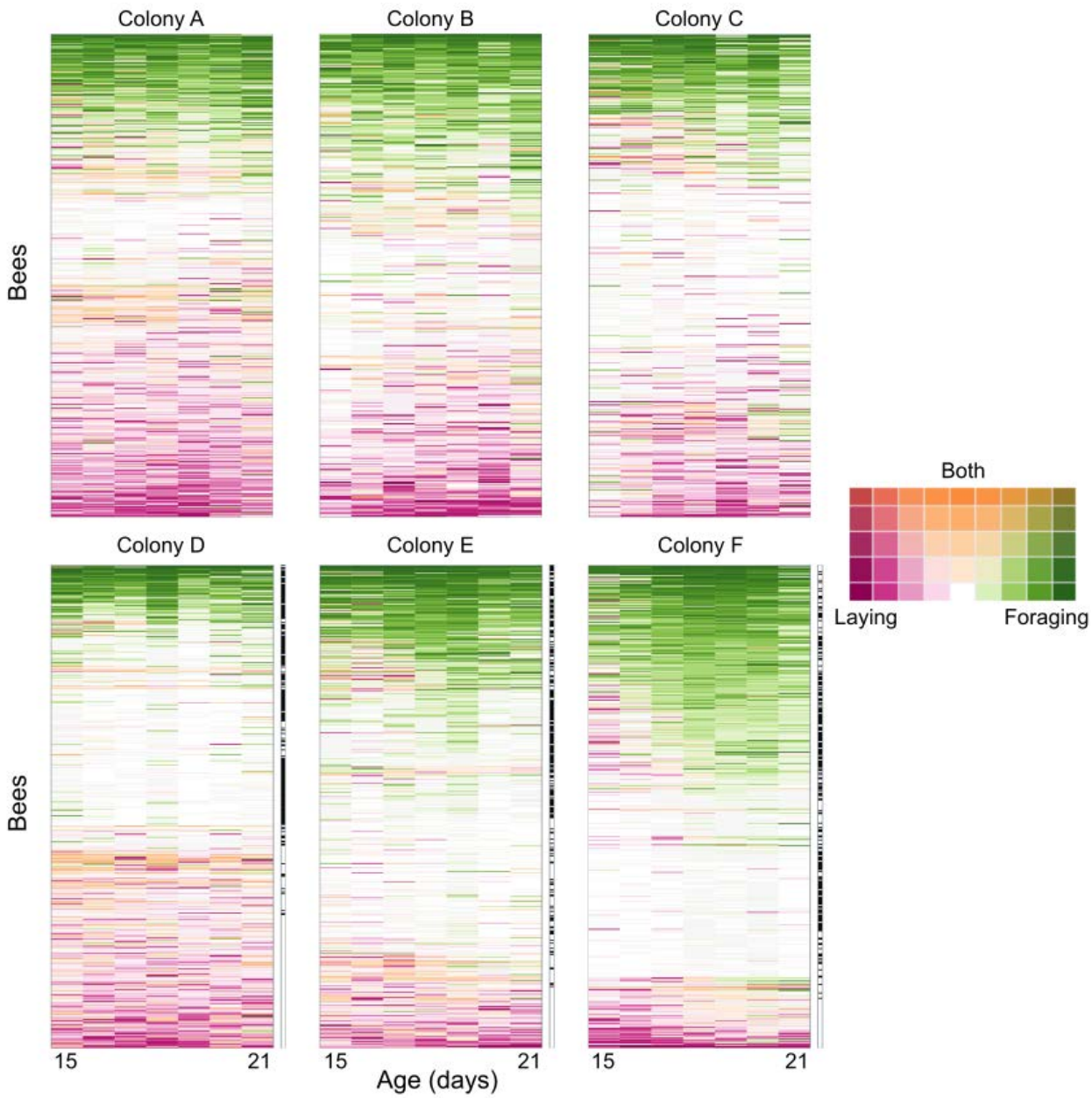


Figure 5.4. Color-mapped behavioral scores for all bees alive at the end of day 21 in each colony. Each row is one bee, and each column is one day. Bees are sorted per colony by median lifetime specialization score, from high (top, specialized on foraging) to low (bottom, specialized on laying). Black and white bars to the right of heatmaps for colonies D-F show source colony of each bee (Colonies D and E: black=SDI 1, white=SDI 2; Colony F: black=SDI 3, white=SDI 4).

Figure 5.5

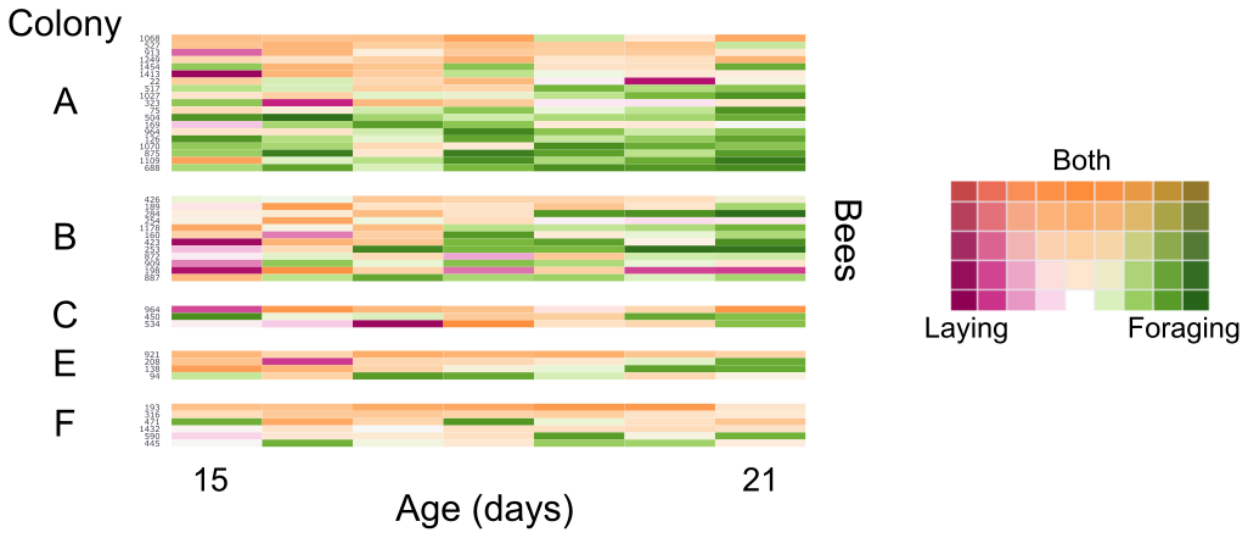


Figure 5.5. Color-mapped behavioral scores for extreme generalists (bees who were classified as generalists for at least 3 days of the experiment using daily behavioral thresholding). Note that some generalists, despite performing both egg-laying and foraging within the same day on multiple occasions, were still highly ranked layers or foragers on other days (dark purple and dark green rectangles).

Figure 5.6

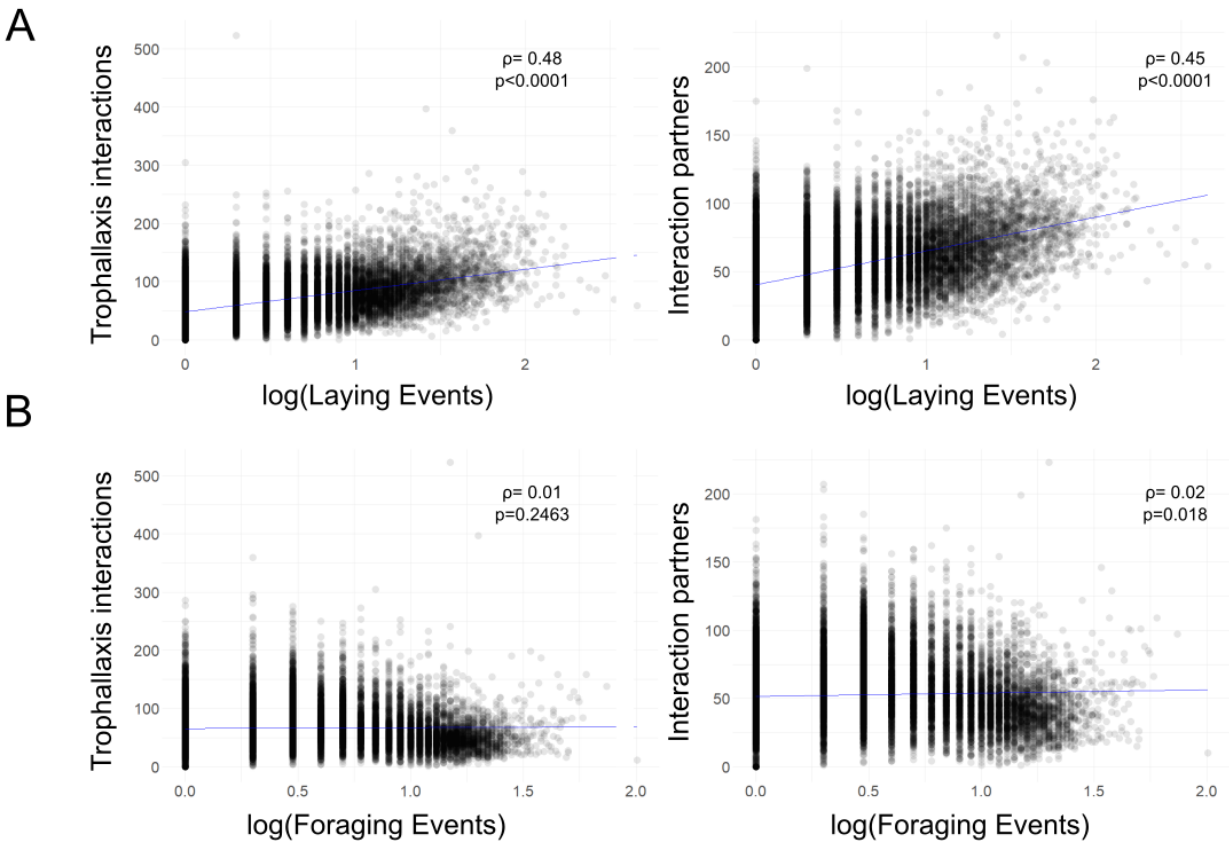


Figure 5.6. Correlations between number of trophallaxis interactions and number of interaction partners with respect to (A) egg-laying and (B) foraging events. Each point represents the daily counts for one individual. One count was added to laying and foraging events prior to log10 transformation. Blue lines indicate regression line, ρ and p-values from Spearman's rank correlation test.

Figure 5.7

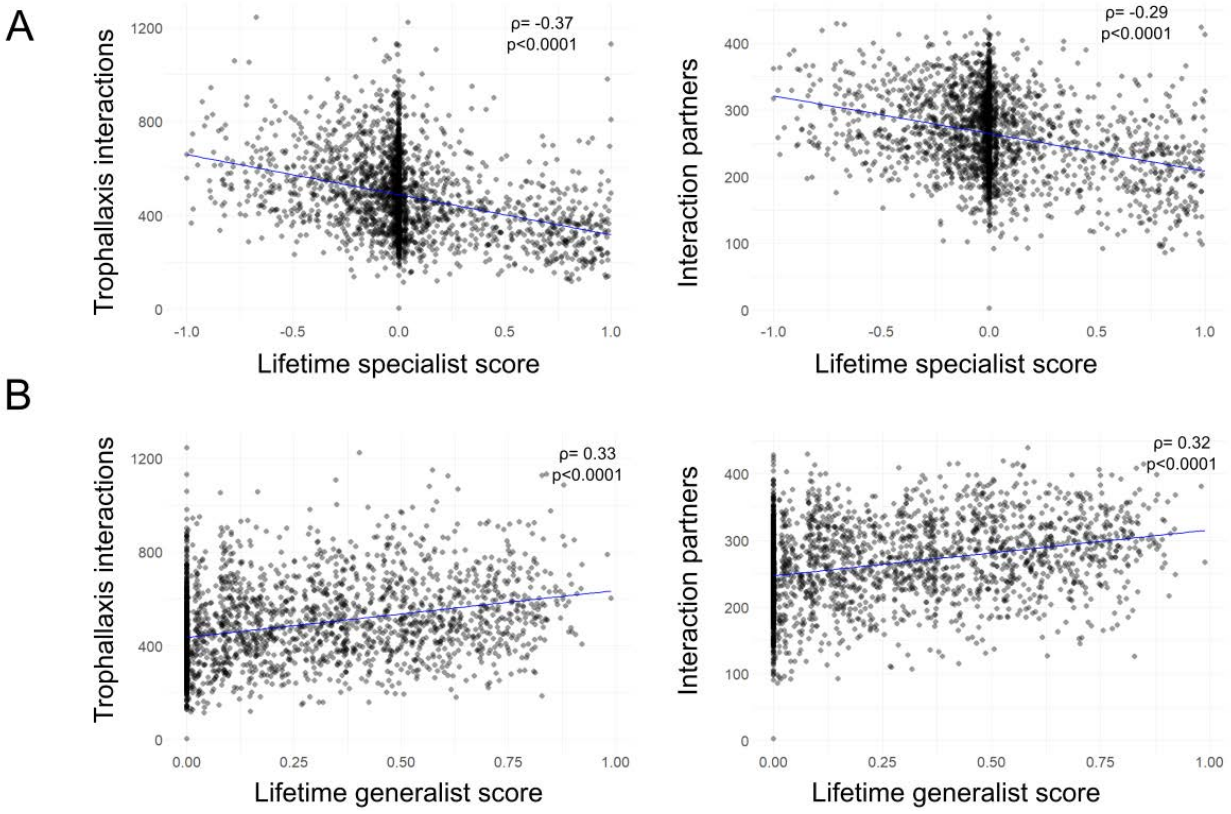


Figure 5.7. Correlations between number of trophallaxis interactions and number of interaction partners with respect to (A) lifetime specialist score (more negative values indicate specialization for egg-laying, while more positive values indicate specialization for foraging) and (B) lifetime generalist score (larger scores indicate great egg-laying and foraging activity). Each point represents one individual. Blue lines indicate regression line, ρ and p-values from Spearman's rank correlation test.

Figure 5.8

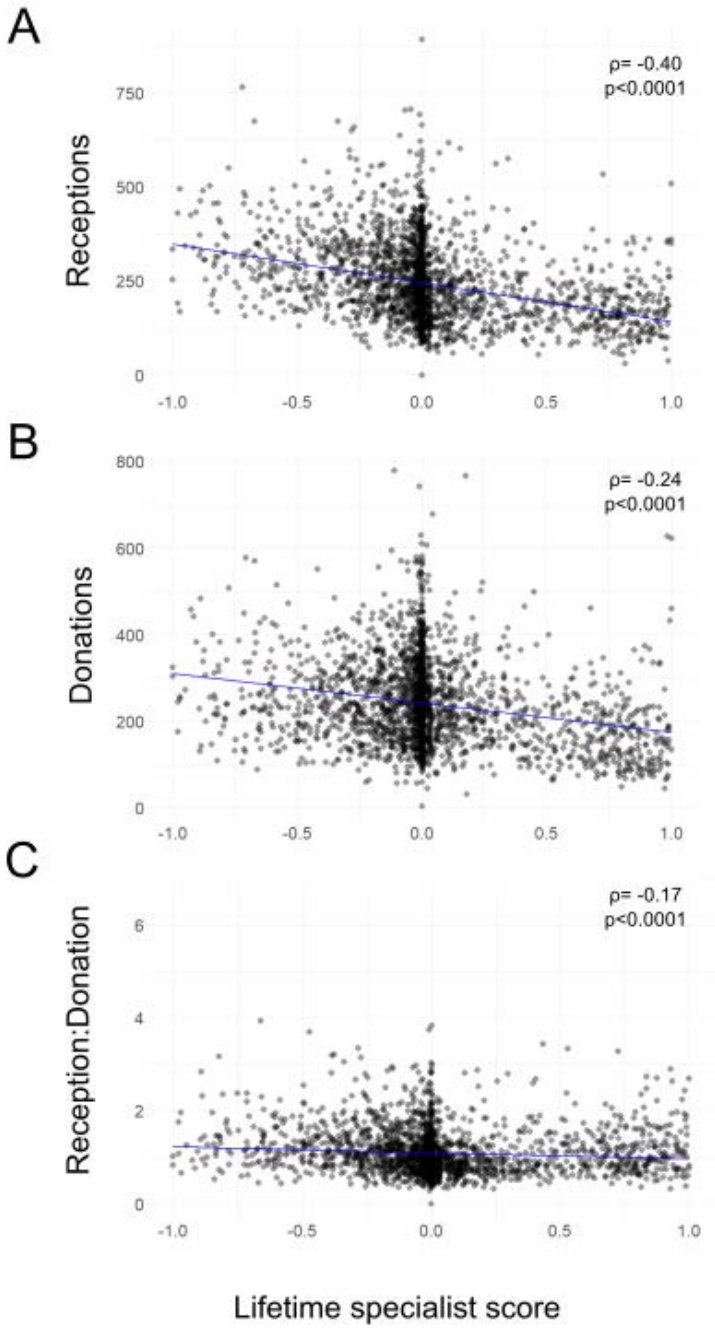


Figure 5.8. Lifetime number of trophallaxis (A) receptions and (B) donations for each bee, as well as the reception:donation ratio (C) plotted against lifetime specialist score (more negative values indicate specialization for egg-laying, while more positive values indicate specialization for foraging). Each point represents one individual. Blue lines indicate regression line, ρ and p-values from Spearman's rank correlation test.

Table 5.1. Rates of egg-laying, foraging, and trophallaxis for each colony.

Behavior	Colony	Rate (counts/bee/day)							
	Day:	15	16	17	18	19	20	21	Mean ± S.E.
Egg-laying (eggs/day)	A	7.4	7.0	6.9	7.0	5.4	6.1	5.6	5.9 ± 0.45
	B	4.3	5.1	5.4	6.7	6.2	5.5	3.9	
	C	4.3	5.9	7.0	5.6	4.4	4.2	1.7	
	D	15.1	15.3	11.5	9.5	7.2	6.2	5.1	
	E	9.2	8.6	6.6	4.4	3.8	3.8	2.3	
	F	5.8	4.5	4.2	3.5	2.6	2.5	1.9	
Foraging (trips/day)	A	5.0	5.4	4.1	4.9	3.2	4.1	4.4	4.0 ± 0.21
	B	4.8	7.2	5.6	4.4	5.6	2.5	6.6	
	C	2.9	3.0	2.7	3.7	2.4	3.1	2.5	
	D	3.9	3.2	2.3	2.9	1.4	1.9	2.0	
	E	4.9	2.9	4.5	5.0	5.6	2.8	3.2	
	F	3.6	4.3	5.1	5.8	6.1	5.6	4.2	
Trophallaxis (interactions/ day)	A	35.8	31.3	24.4	24.5	28.5	31.0	23.6	30.9 ± 1.12
	B	31.4	29.7	25.3	21.8	25.0	22.1	17.9	
	C	42.8	46.8	45.3	40.5	32.1	32.2	19.5	
	D	45.0	46.0	38.3	33.8	28.5	30.5	30.2	
	E	38.4	33.6	33.2	30.0	29.3	26.4	28.2	
	F	33.5	31.4	30.9	27.6	25.6	24.6	21.0	

Table 5.2. Differences in egg-laying and foraging inequality among bees across days, indicating extreme skews within colonies. Values indicate the percentage of top-ranked bees that collectively performed 50% of the total behavioral events of the colony on the given day.

Behavior	Colony	% Bees = 50% Activity							
	Day:	15	16	17	18	19	20	21	Mean ± S.E.
Egg-laying	A	15.3	17.8	17.2	16.1	15.5	11.4	11.6	8.4 ± 0.6
	B	3.4	7.5	8.0	6.5	8.0	8.5	8.3	
	C	4.1	9.2	8.7	8.8	6.6	5.2	2.4	
	D	11.2	11.7	10.4	9.3	9.3	8.6	8.4	
	E	9.5	10.0	9.5	5.9	4.3	3.3	4.8	
	F	12.0	7.6	6.2	4.2	3.0	2.3	1.8	
Foraging	A	11.7	13.5	11.4	13.3	10.6	12.3	12.5	12.1 ± 0.5
	B	12.8	16.9	13.3	15.7	13.7	8.4	13.7	
	C	7.6	10.9	7.2	9.7	7.7	11.1	7.4	
	D	13.5	11.6	8.1	13.9	7.3	9.2	10.5	
	E	12.8	8.5	12.2	12.4	10.6	9.2	8.5	
	F	11.3	15.3	18.1	19.4	18.5	18.3	16.7	

References

Abadi, M., Barham, P., Chen, J., Chen, Z., Davis, A., Dean, J., Devin, M., Ghemawat, S., Irving, G., Isard, M., et al. (2016). TensorFlow : A system for large-scale machine learning. *OSDI* 16, 265–283.

Abrams, J. and Eickwort, G. C. (1981). Nest switching and guarding by the communal sweat bee *Agapostemon virescens* (Hymenoptera, Halictidae). *Insectes Soc.* 28, 105–116.

Amdam, G. V, Csondes, A., Fondrk, M. K. and Page Jr, R. E. (2006). Complex social behaviour derived from maternal reproductive traits. *Nature* 439, 76-78.

Amdam, G. V, Page Jr, R. E. and Tsuruda, J. M. (2008). Sensory response system of social behavior tied to female reproductive traits. *PLoS One* 3, e3397.

Beshers, S. N. and Fewell, J. H. (2001). Models of division of labor in social insects. *Annu. Rev. Entomol.* 46, 413–440.

Beshers, S. N., Huang, Z. Y., Oono, Y. and Robinson, G. E. (2001). Social inhibition and the regulation of temporal polyethism in honey bees. *J. Theor. Biol.* 213, 461–479.

Beverly, B. D., McLendon, H., Nacu, S., Holmes, S. and Gordon, D. M. (2009). How site fidelity leads to individual differences in the foraging activity of harvester ants. *Behav. Ecol.* 20, 633–638.

Bourke, A. F. G. (1988). Dominance orders, worker reproduction, and queen-worker conflict in the slave-making ant *Harpagoxenus sublaevis*. *Behav. Ecol. Sociobiol.* 23, 323–333.

Bourke, A. F. G. (2011). *Principles of Social Evolution*. Oxford: Oxford University Press.

Branstetter, M. G., Danforth, B. N., Pitts, J. P., Faircloth, B. C., Ward, P. S., Buffington, M. L., Gates, M. W., Kula, R. R. and Brady, S. G. (2017). Phylogenomic insights into the evolution of stinging wasps and the origins of ants and bees. *Curr. Biol.* 27, 1019–1025.

Chittka, L. and Muller, H. (2009). Learning, specialization, efficiency and task allocation in social insects. *Commun. Integr. Biol.* 2, 151–154.

Cnaani, J., Schmid-Hempel, R. and Schmidt, J. O. (2002). Colony development, larval development and worker reproduction in *Bombus impatiens* Cresson. *Insectes Soc.* 49, 164–170.

Corona, M., Libbrecht, R., Wurm, Y., Riba-Grognuz, O., Studer, R. A. and Keller, L. (2013). Vitellogenin underwent subfunctionalization to acquire caste and behavioral specific expression in the harvester ant *Pogonomyrmex barbatus*. *PLoS Genet.* 9, e1003730.

Cruz, Y. P. (1981). A sterile defender morph in a polyembryonic hymenopterous parasite. *Nature* 294, 446–447.

Dadant & Sons (1975). Activities of queenless bees. In *The Hive and the Honey Bee* (ed. Dadant & Sons), pp. 245–247. Carthage, Illinois: Journal Printing Company.

Danforth, B. N. (1989). Nesting behavior of four species of *Perdita* (Hymenoptera: Andrenidae). *J. Kansas Entomol. Soc.* 62, 59–79.

Donnell, S. O. (1998). Reproductive caste determination in eusocial wasps (Hymenoptera: Vespidae). *Annu. Rev. Entomol.* 43, 323–346.

Elango, N., Hunt, B. G., Goodisman, M. A. D. and Yi, S. V. (2009). DNA methylation is widespread and associated with differential gene expression in castes of the honeybee, *Apis mellifera*. *Proc. Natl. Acad. Sci.* 106, 11206–11211.

Evans, H. and West-Eberhard, M. (1970). *The Wasps*. Ann Arbor: University of Michigan Press.

Feldmeyer, B., Elsner, D. and Foitzik, S. (2014). Gene expression patterns associated with caste and reproductive status in ants: Worker-specific genes are more derived than queen-specific ones. *Mol. Ecol.* 23, 151–161.

Fine, J. D., Shpigler, H. Y., Ray, A. M., Beach, N. J., Sankey, A. L., Cash-Ahmed, A., Huang, Z. Y., Astrauskaite, I., Chao, R., Zhao, H., et al. (2018). Quantifying the effects of pollen nutrition on honey bee queen egg laying with a new laboratory system. *PLoS One* 13, e0203444.

Free, J. B. (1955). The behaviour of egg-laying workers of bumblebee colonies. *Br. J. Anim. Behav.* 3, 147-153.

Free, J. (1956). A study of the stimuli which release the food begging and offering responses of worker honeybees. *Br. J. Anim. Behav.* 4, 95–101.

Gençer, H. V. and Kahya, Y. (2011). Are sperm traits of drones (*Apis mellifera* L.) from laying worker colonies noteworthy? *J. Apic. Res.* 50, 130–137.

Gernat, T., Rao, V. D., Middendorf, M., Dankowicz, H., Goldenfeld, N. and Robinson, G. E. (2018). Automated monitoring of behavior reveals bursty interaction patterns and rapid spreading dynamics in honeybee social networks. *Proc. Natl. Acad. Sci.* 115, 1433–1438.

Gibbs, J., Brady, S. G., Kanda, K. and Danforth, B. N. (2012). Phylogeny of halictine bees supports a shared origin of eusociality for *Halictus* and *Lasioglossum* (Apoidea: Anthophila: Halictidae). *Mol. Phylogenet. Evol.* 65, 926-939.

Graham, A. M., Munday, M. D., Kaftanoglu, O., Page, R. E., Amdam, G. V. and Rueppell, O. (2011). Support for the reproductive ground plan hypothesis of social evolution and major QTL for ovary traits of Africanized worker honey bees (*Apis mellifera* L.). *BMC Evol. Biol.* 11, 95.

Grozinger, C. M., Fan, Y., Hoover, S. E. R. and Winston, M. L. (2007). Genome-wide analysis reveals differences in brain gene expression patterns associated with caste and reproductive status in honey bees (*Apis mellifera*). *Mol. Ecol.* 16, 4837–4848.

Harrison, M. C., Hammond, R. L. and Mallon, E. B. (2015). Reproductive workers show queenlike gene expression in an intermediately eusocial insect, the buff-tailed bumble bee *Bombus terrestris*. *Mol. Ecol.* 24, 3043–3063.

He, X. J., Jiang, W. J., Zhou, M., Barron, A. B. and Zeng, Z. J. (2017). A comparison of honeybee (*Apis mellifera*) queen, worker and drone larvae by RNA-Seq. *Insect Sci.* 00, 1–11.

Hess, G. (1942). Über den einfluss der weisellosigkeit und des fruchtbarkeitsvitamins e auf die ovarien der bienenarbeiterin. *Schweiz. Bienen Ztg.* 2, 33–110.

Hines, H. M., Hunt, J. H., O'Connor, T. K., Gillespie, J. J. and Cameron, S. A. (2007). Multigene phylogeny reveals eusociality evolved twice in vespid wasps. *Proc. Natl. Acad. Sci.* 104, 3295–3299.

Holldobler, B. and Wilson, E. O. (1990). *The Ants*. Cambridge, MA: Belknap/Harvard University Press.

Hoover, S. E. R., Keeling, C. I., Winston, M. L. and Slessor, K. N. (2003). The effect of queen pheromones on worker honey bee ovary development. *Naturwissenschaften* 90, 477–480.

Hughes, W. O. H., Oldroyd, B. P., Beekman, M. and Ratnieks, F. L. W. (2008). Ancestral monogamy shows kin selection is key to the evolution of eusociality. *Science*. 320, 1213-6.

Hunt, B. G., Wyder, S., Elango, N., Werren, J. H., Zdobnov, E. M., Yi, S. V and Goodisman, M. A. D. (2010). Sociality is linked to rates of protein evolution in a highly social insect. *Mol. Biol. Evol.* 27, 497–500.

Hunt, B. G., Ometto, L., Wurm, Y., Shoemaker, D., Yi, S. V., Keller, L. and Goodisman, M. A. D. (2011). Relaxed selection is a precursor to the evolution of phenotypic plasticity. *Proc. Natl. Acad. Sci.* 108, 15936–15941.

Ihle, K. E., Page, R. E., Frederick, K., Fondrk, M. K. and Amdam, G. V. (2010). Genotype effect on regulation of behaviour by vitellogenin supports reproductive origin of honeybee foraging bias. *Anim. Behav.* 79, 1001–1006.

Jones, B. M., Kingwell, C. J., Wcislo, W. T. and Robinson, G. E. (2017). Caste-biased gene expression in a facultatively eusocial bee suggests a role for genetic accommodation in the evolution of eusociality. *Proc. R. Soc. B Biol. Sci.* 284, 20162228.

Kapheim, K. M. and Johnson, M. M. (2017). Support for the reproductive ground plan hypothesis in a solitary bee: Links between sucrose response and reproductive status. *Proc. R. Soc. B Biol. Sci.* 284, 20162406.

Korst, P. J. A. M. and Velthuis, H. H. W. (1982). The nature of trophallaxis in honeybees. *Insectes Soc.* 29, 209–221.

Kucharski, R., Maleszka, J., Foret, S. and Maleszka, R. (2008). Nutritional control of reproductive status in honeybees via DNA methylation. *Science* 319, 1827–1830.

Kukuk, P. F. and Crozier, R. H. (1990). Trophallaxis in a communal halictine bee *Lasioglossum* (Chilalictus) *erythrurum*. *Proc. Natl. Acad. Sci.* 87, 5402–5404.

Leboeuf, A. C., Benton, R. and Keller, L. (2012). The molecular basis of social behavior: models, methods and advances. *Curr. Opin. Neurobiol.* 23, 3–10.

Leboeuf, A. C., Waridel, P., Brent, C. S., Gonçalves, A. N., Menin, L., Ortiz, D., Riba-Grognuz, O., Koto, A., Soares, Z. G., Privman, E., et al. (2016). Oral transfer of chemical cues, growth proteins and hormones in social insects. *eLife* 5, e20375.

Leichty, A. R., Pfennig, D. W., Jones, C. D. and Pfennig, K. S. (2012). Relaxed genetic constraint is ancestral to the evolution of phenotypic plasticity. *Integr. Comp. Biol.* 52, 16–30.

Leoncini, I., Crauser, D., Robinson, G. E. and Le Conte, Y. (2004). Worker-worker inhibition of honey bee behavioural development independent of queen and brood. *Insectes Soc.* 51, 392–394.

Libbrecht, R., Oxley, P. R. and Kronauer, D. J. C. (2018). Clonal raider ant brain transcriptomics identifies candidate molecular mechanisms for reproductive division of labor. *BMC Biol.* 16, 89.

Liebig, J., Heinze, J. and Hölldobler, B. (1997). Trophallaxis and aggression in the ponerine ant, *Ponera coarctata*: implications for the evolution of liquid food exchange in the Hymenoptera. *Ethology* 103, 707–722.

Linksvayer, T. A. and Wade, M. J. (2005). The evolutionary origin and elaboration of sociality in the aculeate Hymenoptera: maternal effects, sib-social effects, and heterochrony. *Q. Rev. Biol.* 80, 317–336.

Malka, O., Shnieor, S., Katzav-Gozansky, T. and Hefetz, A. (2008). Aggressive reproductive competition among hopelessly queenless honeybee workers triggered by pheromone signaling. *Naturwissenschaften* 95, 553–559.

McCorquodale, D. B. (1989). Soil softness, nest initiation and nest sharing in the wasp, *Cerceris antipodes* (Hymenoptera: Sphecidae). *Ecol. Entomol.* 14, 191–196.

McCorquodale, D. B. and Naumann, I. D. (1988). A new Australian species of communal ground nesting wasp, in the genus *Spilomena* Shuckard (Hymenoptera: Sphecidae: Pemphredoninae). *Aust. J. Entomol.* 27, 221–232.

Michener, C. D. (1974). *The Social Behavior of the Bees: A Comparative Study*. Cambridge, MA: Harvard University Press.

Miller III, D. G. and Ratnieks, F. L. W. (2001). The timing of worker reproduction and breakdown of policing behaviour in queenless honey bee (*Apis mellifera* L.) societies. *Insectes Soc.* 48, 178.

Mohammedi, A., Paris, A., Crauser, D. and Le Conte, Y. (1998). Effect of aliphatic esters on ovary development of queenless bees (*Apis melliera* L.). *Naturwissenschaften* 85, 455–458.

Moreau, C. S., Bell, C. D., Vila, R., Archibald, S. B. and Pierce, N. E. (2006). Phylogeny of the ants: Diversification in the age of angiosperms. *Science* 312, 101–104.

Morse, R. A. (1990). Laying workers. In *The ABC & XYZ of Bee Culture* (eds. Morse, R. A. and Flottum, K.), pp. 291–293. The A. I. Root Company.

Naeger, N. L., Peso, M., Even, N., Barron, A. B. and Robinson, G. E. (2013). Altruistic behavior by egg-laying worker honeybees. *Curr. Biol.* 23, 1574–1578.

Nixon, H. L. and Ribbands, C. R. (1952). Food transmission within the honeybee community. *Proc. R. Soc. B Biol. Sci.* 140, 43–50.

O'Donnell, S. and Jeanne, R. L. (1990). Forager specialization and the control of nest repair in *Polybia occidentalis* Olivier (Hymenoptera : Vespidae). *Behav. Ecol. Sociobiol.* 27, 359–364.

Oster, G. F. and Wilson, E. O. (1979). *Caste and Ecology in the Social Insects*. Princeton: Princeton University Press.

Page Jr, R. E. and Robinson, G. E. (1994). Reproductive competition in queenless honey bee colonies (*Apis mellifera* L.). *Behav. Ecol. Sociobiol.* 35, 99–107.

Page Jr, R. E. and Amdam, G. V (2007). The making of a social insect: developmental architectures of social design. *Bioessays* 29, 334.

Page Jr, R. E. and Erickson E H, J. (1988). Reproduction by worker honey bees (*Apis mellifera* L.). *Behav. Ecol. Sociobiol.* 23, 117–126.

Page Jr, R. E., Scheiner, R., Erber, J. and Amdam, G. V (2006). The development and evolution of division of labor and foraging specialization in a social insect (*Apis mellifera* L.). *Curr. Top. Dev. Biol.* 74, 253–286.

Pamminger, T. and Hughes, W. O. H. (2017). Testing the reproductive groundplan hypothesis in ants (Hymenoptera: Formicidae). *Evolution* 71, 153–159.

Pardi, L. (1948). Dominance order in *Polistes* wasps. *Physiol. Zool.* 21, 1–13.

Pruitt, J. N. and Riechert, S. E. (2011). How within-group behavioural variation and task efficiency enhance fitness in a social group. *Proc. R. Soc. B Biol. Sci.* 278, 1209–1215.

Ratnieks, F. L. W. (1993). Egg-laying, egg-removal, and ovary development by workers in queenright honey bee colonies. *Behav. Ecol. Sociobiol.* 32, 191–198.

Ratnieks, F. L. W. and Wenseleers, T. (2008). Altruism in insect societies and beyond: voluntary or enforced? *Trends Ecol. Evol.* 23, 45–52.

Ratnieks, F. L. W., Foster, K. R. and Wenseleers, T. (2006). Conflict resolution in insect societies. *Annu. Rev. Entomol.* 51, 581–608.

Rehan, S. M. and Toth, A. L. (2015). Climbing the social ladder: the molecular evolution of sociality. *Trends Ecol. Evol.* 30, 426–433.

Ribbands, C. R. (1952). Division of labour in the honeybee community. *Proc. R. Soc. B. Biol. Sci.* 140, 32–43.

Robinson, G. E. (1985). Effects of a juvenile hormone analogue on honey bee foraging behaviour and alarm pheromone production. *J. Insect Physiol.* 31, 277–282.

Robinson, G. E., Page, R. E. and Fondrk, M. K. (1990). Intracolonial behavioral variation in worker oviposition, oophagy, and larval care in queenless honey bee colonies. *Behav. Ecol. Sociobiol.* 26, 315–323.

Romiguier, J., Cameron, S. A., Woodard, S. H., Fischman, B. J., Keller, L. and Praz, C. J. (2016). Phylogenomics controlling for base compositional bias reveals a single origin of eusociality in corbiculate bees. *Mol. Biol. Evol.* 33, 670–678.

Sakagami, S. F. (1954). Occurrence of an aggressive behaviour in queenless hives, with considerations on the social organization of honeybee. *Insectes Soc.* 1, 331–343.

Sakagami, S. F. (1958). The False-Queen: fourth adjustive response in dequeened honeybee colonies. *Behaviour* 13, 280–295.

Schindelin, J., Arganda-Carreras, I., Frise, E., Kaynig, V., Longair, M., Pietzsch, T., Preibisch, S., Rueden, C., Saalfeld, S., Schmid, B., et al. (2012). Fiji: An open-source platform for biological-image analysis. *Nat. Methods* 9, 676-682.

Schwarz, M. P., Richards, M. H. and Danforth, B. N. (2007). Changing paradigms in insect social evolution: insights from halictine and allodapine bees. *Annu. Rev. Entomol.* 52, 127–150.

Sekiguchi, K. and Sakagami, S. F. (1966). Structure of foraging population and related problems in the honey bee, with considerations on the division of labour in bee colonies. *Hokkaido Nat. Agric. Exp. Sta. Rep.* 69, 1–65.

Shi, Y. Y., Liu, H., Qiu, Y. F., Ma, Z. Y. and Zeng, Z. J. (2017). DNA methylation comparison between 4-day-old queen and worker larvae of honey bee. *J. Asia. Pac. Entomol.* 20, 299–303.

Simola, D. F., Graham, R. J., Brady, C. M., Enzmann, B. L., Desplan, C., Ray, A., Zwiebel, L. J., Bonasio, R., Reinberg, D., Liebig, J., et al. (2015). Epigenetic (re)programming of caste-specific behavior in the ant *Camponotus floridanus*. *Science* 351, aac6633.

Sommeijer, M. J. and Van Veen, J. W. (1990). The polygyny of *Myrmica rubra*: selective oophagy and trophallaxis as mechanisms of reproductive dominance. *Entomol. Exp. Appl.* 56, 229–239.

Tenczar, P., Lutz, C. C., Rao, V. D., Goldenfeld, N. and Robinson, G. E. (2014). Automated monitoring reveals extreme interindividual variation and plasticity in honeybee foraging activity levels. *Anim. Behav.* 95, 41–48.

Toth, A. L., Varala, K., Newman, T. C., Miguez, F. E., Hutchison, S. K., Willoughby, D. A., Simons, J. F., Egholm, M., Hunt, J. H., Hudson, M. E., et al. (2007). Wasp gene expression supports an evolutionary link between maternal behavior and eusociality. *Science* 318, 441–444.

Trumbo, S. T. and Robinson, G. E. (1997). Learning and task interference by corpse-removal specialists in honey bee colonies. *Ethology* 103, 966–975.

Turillazzi, S. and West-Eberhard, M. (1996). *Natural History and Evolution of Paper-Wasps*. Oxford: Oxford University Press.

Velthuis, H. H. W. (1970). Ovarian development in *Apis mellifera* worker bees. *Entomol. Exp. Appl.* 13, 377-394.

Velthuis, H. H. W. and Duchateau, M. J. (2011). Ovarian development and egg laying in workers of *Bombus terrestris*. *Entomol. Exp. Appl.* 51, 199–213.

Visscher, P. K. (1989). A quantitative study of worker reproduction in honey bee colonies. *Behav. Ecol. Sociobiol.* 25, 247–254.

Visscher, P. K. (1996). Reproductive conflict in honey bees: A stalemate of worker egg-laying and policing. *Behav. Ecol. Sociobiol.* 39, 237–244.

Vollbehr, V. J. (1975). The orientation of young honey bees on their first orientation flight. *Zool. Jb. Physiol. Bd.* 79, 33–69.

Walton, A. and Toth, A. L. (2016). Variation in individual worker honey bee behavior shows hallmarks of personality. *Behav. Ecol. Sociobiol.* 70, 999–1010.

Wenseleers, T. and Ratnieks, F. L. W. (2006). Enforced altruism in insect societies. *Nature* 444, 50.

West-Eberhard, M. J. (1987). *Flexible strategy and social evolution*. (eds. Ito, Y., Brown, J. L., and Kikkawa, J.). Tokyo: Japan Scientific Societies Press.

West-Eberhard, M. J. (1996). Wasp societies as microcosms for the study of development and evolution. In *Natural history and evolution of paper wasps (eds Turillazzi S, West-Eberhard MJ)*, pp. 290–317.

Wheeler, D. E. (1986). Developmental and physiological determinants of caste in social Hymenoptera: evolutionary implications. *Am. Nat.* 128, 13–34.

Wilson, E. O. (1971). *The insect societies*. Cambridge: Harvard University Press.

Wilson, E. O. (1980). Caste and division of labor in leaf-cutter ants (Hymenoptera: Formicidae: Atta). *Behav. Ecol. Sociobiol.* 7, 157–165.

Wilson, E. O. (1985). The sociogenesis of insect colonies. *Science* 228, 1489–95.

Winston, M. L. (1987). *The Biology of the Honey Bee*. Cambridge, MA: Harvard University Press.

Winston, M. L. and Katz, S. J. (1982). Foraging differences between cross-fostered honeybee workers (*Apis mellifera*) of European and Africanized races. *Behav. Ecol. Sociobiol.* 10, 125–129.

CHAPTER 6

SOCIAL PLASTICITY IN LAYING WORKER HONEY BEES ENABLED BY CONTINUOUS VARIATION IN GENE EXPRESSION AND CHROMATIN LANDSCAPES

Abstract

Understanding the origin of phenotypes from ancestral gene regulatory networks is a fundamental goal of evolutionary biology. Eusociality represents an extreme form of social behavior which has evolved independently multiple times across insects and is characterized by a novel polyphenism that gives rise to reproductive (queen) and non-reproductive (worker) castes. Many hypotheses related to the origin of eusocial castes implicate gene regulatory changes, but how transitions between social forms may be enabled by plasticity in gene regulatory networks is unknown. In Chapter 5, I described the surprising presence of multiple levels of social plasticity in queenless laying worker (LW) colonies of the obligate and advanced eusocial honey bee, *Apis mellifera.* This included a number of "generalist" bees who engage in reproductive and non-reproductive behaviors much like solitary or communal species, possible ancestral social states of the honey bee lineage, as well as workers who engaged in a reproductive division of labor similar to primitively eusocial societies. In this study, I utilized the socio-behavioral plasticity present in LW colonies to explore how flexibility in gene expression and gene regulation may contribute to variation in social organization. I found that LW honey bees specialized at either end of a reproductive division of labor (egg-laying or foraging) displayed widespread differences in brain gene expression, with corresponding differences in accessible chromatin. Forager-biased genes showed a strong signature of metabolic activity, including upregulation of the glycolysis pathway, and were enriched for many caste-related genes and genes under selection in other species of social insects. Egg-laying workers had brain gene expression profiles similar to those of queens in two species with facultative or primitive eusociality, and were also enriched for genes upregulated in "nurse" (brood care providing) workers in queenright honey bee colonies. This result is consistent with previously published hypotheses on the co-option of ancestral solitary reproductive signaling pathways during the evolution of derived, specialized honey bee subcastes. My findings on generalist bees, who engaged in both reproductive and non-

reproductive behaviors, provide insights into how this co-option might have occurred in evolution. Generalists displayed intermediate transcriptomic and accessible chromatin profiles relative to either specialized group, suggesting a surprising degree of fluidity between reproductive and non-reproductive phenotypes despite the highly derived eusocial status of the honey bee. Together, these results suggest that variation in social organization may emerge through latent, environmentally-induced incremental changes in the regulation of conserved pathways.

Introduction

Phenotypic plasticity is increasingly being recognized as an important influencer of evolution (Moczek et al., 2011; Pfennig et al., 2010; Pigliucci, 2006; West-Eberhard, 2003), but how that plasticity emerges from ancestral gene regulatory networks is still an area of active interest. The evolution of some plastic phenotypes, such as many morphological traits (Abzhanov et al., 2006; Emlen, 2000; Emlen and Nijhout, 2000), has been elucidated through the integration of developmental genetics with comparative evolutionary biology in the field of evo-devo (Carroll, 2008; Raff, 2000). Other plastic phenotypes, including complex traits such as social behavior, are more challenging to study from both mechanistic as well as evolutionary perspectives. These challenges include extensive gene by environment interactions (Zayed and Robinson, 2012), the influence of prior experience on future behavior (Dall et al., 2005; Shpigler et al., 2017b), and interactions among individuals.

The most extreme form of social behavior, eusociality, is defined by an overlap of adult generations, cooperative care of brood, and a reproductive division of labor among queen and worker castes (Wilson, 1971). Eusociality represents an evolutionary transition which has been likened to the transition from unicellularity to multicellularity (Smith and Szathmaáry, 1995), as individual members of eusocial colonies cooperate within a reproducing “superorganism” (Hölldobler and Wilson, 2008). Much like specialized cells and tissues emerge from totipotent stem cells, queen and worker castes of eusocial species typically arise from environmentally-mediated bifurcating developmental pathways. In obligate and advanced eusocial species, queens and workers are highly specialized for reproductive and non-reproductive functions, respectively, with multiple differentiating morphological, physiological, and behavioral features despite a common genomic background. How these castes evolved from a solitary ancestor is of interest

not only for the understanding of social evolution, but also important for understanding how ancestral gene regulatory networks can be modified for increased phenotypic plasticity and complexity.

Many studies of molecular caste evolution have utilized a comparative approach across the social insects, where eusociality has evolved independently at least 9 times (Bourke, 2011; Cruz, 1981; Gibbs et al., 2012; Hines et al., 2007; Hughes et al., 2008; Moreau et al., 2006; Romiguier et al., 2016; Schwarz et al., 2007). These studies compare extant snapshots of different social organizations across species and identify common genomic features associated with level of sociality (Kapheim et al., 2015; Woodard et al., 2011). By comparing across species, however, these studies are constrained by the unique evolutionary histories and pressures experienced by different lineages. Groups or species with flexible social strategies are thus also important to study, and have provided additional insights into genomic features of evolution associated with sociality (Jones et al., 2017; Kapheim et al., 2013; Kocher and Paxton, 2014). Many of these groups lack the highly specialized castes of obligate eusocial species, however, which limits understanding of how the extremely specialized castes of advanced eusocial species have evolved.

In Chapter 5, I described the discovery of a new social state for colonies of the obligate and advanced eusocial Western honey bee, *Apis mellifera.* By using high-resolution and automated behavioral tracking, I discovered that queenless, broodless colonies of honey bees, which result in the appearance of egg-laying workers (LW), exhibit high levels of behavioral plasticity both at the individual and colony levels. Colonies with LW exhibit multiple social phenotypes. This includes some individuals that share a nest while expressing behavioral plasticity resembling the life history of extant solitary bees, which is thought to represent the ancestral life history of eusocial bees (Michener, 1974; Wilson, 1971). Other individuals participate in a division of labor for reproductive and non-reproductive behaviors, similar to what is observed in primitively eusocial species of bees and wasps. This unique colony organization, part communal (totipotent individuals engaging in personal reproduction while cohabiting the same nest; Michener, 1974) and part primitively eusocial (totipotent individuals engaged in a reproductive division of labor; Michener, 1974), is a social state previously undescribed in any species of social insect. The presence of these multiple social phenotypes suggests that LW colonies may provide a glimpse into a transitional stage in social evolution, where solitary-like

individuals may have co-occurred with individuals that begin to decouple the traits later associated with distinct queen and worker castes. Molecular studies of LW colonies may therefore provide insights into how castes evolved in eusocial societies, as well as help to better understand the molecular underpinnings of variation in social organization across social insects.

The regulation of behavioral phenotypes, including those modulated by the social environment, is known to involve gene expression changes in the brain. Both short-term transcriptional changes associated with acute stimulus presentation (Bukhari et al., 2017; Saul et al., 2017; Shpigler et al., 2017a), as well as stable differences in gene expression between behavioral states (Whitfield et al., 2003; Zayed and Robinson, 2012), have been described for many species. In social insects, queen and worker castes are marked by differential gene expression throughout development and as adults, in the brain as well as in other tissues (Feldmeyer et al., 2014; Grozinger et al., 2007; Harrison et al., 2015; He et al., 2017; Jones et al., 2017). Within the worker caste, brain transcriptional profiles between nurse and forager worker honey bees are highly distinct (Alaux et al., 2009a; Ament et al., 2012; Whitfield et al., 2003), reflecting the stable differences in behavioral phenotypes between these subcastes. Responses to colony threats (Alaux et al., 2009b; Shpigler et al., 2017a), food rewards (McNeill et al., 2016), and exposure to pheromones (Alaux et al., 2009a; Grozinger et al., 2003) all lead to brain gene expression changes, demonstrating the dynamic transcriptome of the honey bee brain (Zayed and Robinson, 2012).

Eukaryotic transcription is mediated through the differential binding of transcription factors and other protein complexes involved in the transcription apparatus (Latchman, 1993; Orphanides and Reinberg, 2002; Spitz and Furlong, 2012). A number of epigenetic mechanisms influence this binding, including structural changes in chromatin (Kornberg and Lorch, 1995; Narlikar et al., 2002; Venkatesh and Workman, 2015), histone modifications (Bannister and Kouzarides, 2011; Grunstein, 1997), and DNA methylation (Cedar, 1988; Jones and Takai, 2001; Kass et al., 1997). The downstream effects of transcription are additionally regulated by negative feedback mechanisms including transcript degradation through microRNAs (Filipowicz et al., 2008; Hobert, 2008; Huang et al., 2011) and other RNA-binding proteins which can be target-specific or general to all expressed transcripts (Guhaniyogi and Brewer, 2001).

In social insects, many of the above transcriptional regulatory mechanisms have been discovered to influence caste determination. Genome-wide differences in histone modifications

have been described in developing queens and workers (Wojciechowski et al., 2018), and differential expression of microRNAs has been reported in developing queen and worker larvae of the honey bee (Ashby et al., 2016; Shi et al., 2015). Differential methylation has also been implicated in the development of queen and worker castes (Elango et al., 2009; Foret et al., 2012; Kucharski et al., 2008; Lyko et al., 2010; Shi et al., 2017), as well as differences in worker subcastes of the honey bee (Herb et al., 2012). Even short-term exposure to socially-relevant stimuli in the honey bee leads to changes in histone modifications and DNA methylation in the brain and is associated with differential gene expression (Herb et al., 2018; Shpigler et al., 2017a; Shpigler et al., 2018). The relationship between chromatin structure and gene expression is mediated in part through transcription factors, which are predicted to regulate honey bee behavioral states through context-specific transcriptional modules (Chandrasekaran et al., 2011; Khamis et al., 2015).

In this study, I selected representatives of the different behavioral states of queenless LW colonies and sequenced their brain mRNA and accessible chromatin to gain insights into how the unique and evolutionarily relevant social organization of LW colonies may be achieved. Generalists, those bees who express behavioral plasticity similar to solitary bees, as well as bees engaged in a reproductive division of labor (specialized egg-layers and foragers) were selected from two LW colonies studied in Chapter 5 for molecular dissection. If specialized egg-layers and foragers represent stable behavioral subcastes like those seen in queenright colonies (e.g., nurses and foragers), the brain transcriptional profiles of these groups should be distinct and likely regulated through differences in accessible chromatin. Alternatively, this division of labor may be mediated through differences in neuronal activity that occur on a shorter-timescale than changes in gene expression, perhaps reflective of more dysregulated behavioral states. Generalists, who both lay eggs and forage, may show unique brain gene expression and chromatin accessibility profiles relative to either specialized group, reflecting their expanded behavioral plasticity. Alternatively, generalists may fall intermediate along a continuum of gene expression and chromatin accessibility, representing an ancestral “middle ground” relative to specialized groups. This would be reminiscent of the ancestral form that then was subject to disruptive selection in the transition from solitary to group living (Eberhard, 1980; West-Eberhard, 1987). This mapping of social plasticity in molecular terms not only addresses questions related to the regulation of social behavior, but also provides insights into how the

regulation of behavioral plasticity may differ relative to plasticity in other phenotypes.

Methods

Bees

Source colonies

Honey bee colonies were maintained according to standard beekeeping practices at the University of Illinois Bee Research Facility in Urbana, Illinois. Experimental colonies were established from a mix of two source colonies each headed by an artificially inseminated queen, who had been single-drone inseminated (SDI) and was of either *Apis mellifera ligustica* or *Apis mellifera carnica* origin (queen rearing and inseminations performed by Sue Cobey- Honey Bee Insemination Service; Washington State University). One of these SDI colonies had been previously screened using age-matched workers in laboratory cages and demonstrated rapid development of ovaries and egg-laying within 11 days (Appendix Figure A.5). One-day-old adult worker bees from source colonies were obtained by removing sealed frames of late-stage pupae from source colonies and housing them in an incubator inside emergence cages at 34 °C and 50% relative humidity. Bees were swept from frames daily to collect adults less than 24 hours old.

Barcoding bees

Bees were barcoded with custom "bCodes" as in Gernat et al. (2018). Unique sets of bCodes were used to differentiate bees barcoded on different days, as well as to differentiate bees from different source colonies. Briefly, workers were anesthetized on ice and then positioned using soft forceps (BioQuip Products). A small drop of Loctite Super Glue Gel Control (Henkel) was applied to the center of the thorax of each bee, followed by a bCode positioned with its orientation vector parallel to the anteroposterior axis of the bee. Bees were carefully placed in plastic dishes until they recovered from cold anesthetization, at which point the glue was dry. After waking, all bees were placed in a large container with Fluon®-coated walls (Insect-a-Slip, BioQuip) where honey was provided *ab libitum* until placement into the hive. At the end of each barcoding day, bees were carefully transferred into a custom observation hive, described below.

Behavioral tracking

Observation hives and entrances were monitored as in Chapter 5. Briefly, barcoded bees were housed in a glass-walled observation hive with access to the outside through an entrance tunnel. Hive images were acquired continuously at one-second resolution using infrared lights

and a Prosilica GX6600 machine vision camera (Allied Vision). Entrance images were acquired twice per minute with a Raspberry Pi camera (5 megapixel v1.3, Adafruit) from 07:00 to 19:00. Images were saved as JPEGs and copied onto a computing cluster (Biocluster, UIUC) for analysis. Barcode detection and automatic behavioral tracking was performed as in Chapter 5, providing one-second-resolution behavioral data for three behaviors: egg-laying, foraging, and trophallaxis.

At the end of behavioral tracking, observation colonies were placed into a -80°C chest freezer for 30 minutes or until all bees were immobilized. Next, the frame (with bees immobilized) was placed into a large tray of dry ice, and soft forceps were used to individually flash freeze bees in liquid nitrogen before being placed into 96 deep well plates. Plates were stored at -80°C until dissection and tissue preparation, described below.

Selection of bees for sequencing

Each bee was given two scores based on their performance of egg-laying and foraging behavior over time as in Chapter 5. The specialist score describes how specialized an individual was on either egg-laying (scores near -1) or foraging (scores near +1) relative to other bees in the colony; bees that consistently performed both egg-laying and foraging, or that performed neither behavior, have specialist scores near 0. The generalist score ranges from 0 to 1 and describes the degree to which an individual performed both egg-laying and foraging behaviors, differentiating bees with specialist scores near 0 based on the performance (or not) of egg-laying and foraging.. A high generalist score indicates high rank for both egg-laying and foraging on a given day. The median of both scores (weighted to emphasize the latter part of the experiment) across days was used to characterize the overall behavior of each bee in the colony. The rank approach allows for normalization across days with different overall levels of activity in the colony, and the median score across days provides an overall assessment of the lifetime behavior of each bee.

These weighted median scores were used to rank all bees, and the top ranking specialists and generalists were selected for sequencing in both colonies (Fig. 6.1). Scores for each sequenced bee (n=45, 20 from colony R18, 25 from colony R25), as well as total numbers of detected egg-laying and foraging events per bee, are provided in Appendix Table B.3.

Tissue dissection and homogenization

Abdomens of each bee were carefully removed on dry ice and incubated for 16 hours at -20°C in RNA-later ICE (Life Technologies). Ovaries were imaged and assessed for ovary

development using a 1-5 scale adapted from Hess (1942) to confirm that egg-layers and generalists had activated ovaries (3-5 on scale). Ovary scores, as well as number of ovarioles as determined from dissections, are given in Appendix Table B.3.

The heads of each bee were freeze-dried at 300 milliTorr for 55 minutes, and whole brains were removed from the head capsule in a dry ice ethanol bath. Dissected brains were stored individually in 1.5 mL microcentrifuge tubes at -80°C until extractions.

Brains were individually homogenized in 150 uL phosphate buffered saline (1X PBS, Corning, cat # 21-040-CV) with protein inhibitor complex (PIC, Complete Tablets, EDTA-free Protease Inhibitor Cocktail from Roche, cat # 04693132001) using a motorized pestle for 20 seconds. 50 uL of this homogenate was then pipetted into 450 uL cold PBS+PIC and placed on ice for ATAC-seq library preparation (see below). The remaining 100 uL homogenate was mixed with 500 uL RLT buffer (Qiagen) with 1% β-mercaptoethanol for use in the Qiagen RNeasy Mini Kit RNA extraction protocol (see below).

RNAseq library preparation and sequencing

Whole brain RNA was extracted from the 600 uL homogenate in RLT buffer after an additional 30 second homogenization following the Qiagen RNeasy Mini Kit protocol, including a DNase (Qiagen) treatment to remove genomic DNA. RNA quantities were determined for each sample using a Qubit RNA HS Assay Kit (Invitrogen). High RNA integrity for all samples was confirmed with Bioanalyzer 2100 RNA Pico chips (Agilent) prior to library preparation.

RNAseq libraries were constructed and sequenced by the W.M. Keck Center for Comparative and Functional Genomics at the Roy J. Carver Biotechnology Center (University of Illinois at Urbana-Champaign). Libraries were constructed from 500 ng RNA per sample using the TruSeq Stranded mRNA HT kit (Illumina) on an ePMotion 5075 robot (Eppendorf). Libraries were uniquely barcoded, quantified, and pooled for sequencing across 6 lanes with 100 nt single-end sequencing on the Illumina HiSeq 4000.

ATACseq library preparation and sequencing

The 500 uL tissue homogenate was additionally homogenized by aspirating through a 20 gauge needle followed by a 23 gauge needle 5 times each. Samples were centrifuged at 500g for 5 minutes at 4°C. Supernatant was removed, and cells were resuspended in 50 uL cold PBS+PIC. 15 uL of this cell suspension (approximately $1/10^{th}$ of the total brain, ~100k cells) was placed into a new microcentrifuge tube, and this was centrifuged at 500g for 5 minutes at 4°C as an

additional cell washing step. Supernatant was removed, and cells were gently resuspended in 50 uL cold lysis buffer prepared as in Buenrostro et al. (2015). The remainder of the ATACseq library protocol followed Buenrostro et al. (2015), with the exception of the final purification step, where a 0.8:1 ratio of Ampure XP beads to sample (Beckman Coulter) was used to purify each library. In addition to sample libraries, input libraries were constructed from thoracic genomic DNA from a sister of each colony per sequencing batch using 50 ng of genomic DNA (extracted using the Gentra Puregene Tissue Kit from Qiagen, cat # 158667, following manufacturer's protocol for DNA purification from 25 mg tissue but with 6 uL proteinase K and 4 uL RNase A at the appropriate steps). Genomic DNA was transposed following the ATACseq protocol immediately following the cell lysis step (Buenrostro et al. 2015), again using an Ampure XP bead clean-up at the end of the protocol. A Qubit dsDNA HS Assay Kit (Invitrogen) was used to quantify each library, and library size and quality was assessed using a Bioanalyzer High-Sensitivity DNA Analysis kit (Agilent).

ATACseq libraries, including input libraries, were pooled at equal nM concentrations and a bead clean-up (0.8:1 ratio of Ampure XP beads to sample) was performed on the pool prior to submission for sequencing. QC on the final pool was performed using qPCR and an AATI Fragment Analyzer by the W.M. Keck Center for Comparative and Functional Genomics at the Roy J. Carver Biotechnology Center (University of Illinois at Urbana-Champaign). Libraries were sequenced across three lanes with 100 nt paired-end sequencing on the Illumina HiSeq 4000 by the Keck Center.

Data processing and analysis

RNAseq

Sequencing of RNAseq libraries (n=45, 20 from colony A, 25 from colony B) produced 1,487,641,973 reads which survived quality and adapter trimming using Trimmomatic (version 0.36, parameters used: ILLUMINACLIP: 2:35:30 LEADING:20 TRAILING:20 MINLEN:30). Trimmed reads were aligned to the *Apis mellifera* HAv3.1 genome (NCBI accession GCA_003254395.2) using STAR (version 2.5.3) and default parameters, resulting in an average of 96.7% reads mapping uniquely. featureCounts from the Subread package (version 1.5.2) was used to assign mapped reads to gene features from the GFF file from NCBI associated with the *A. mellifera* HAv3.1 genome. On average, 84.8% of uniquely mapped reads were assigned to gene features using featureCounts.

Gene counts were imported into R for differential expression analysis using edgeR. Genes with less than 1 CPM in at least 2 samples were removed, and remaining count values were normalized using the TMM method. Tagwise dispersion estimates were followed by quasi-likelihood F tests for each pairwise comparison of groups. Differential gene expression (DEG) results for each pairwise comparison are given in Appendix Dataset C.1.

ATACseq

Sequencing of ATACseq libraries (n=48, 20 from colony A, 25 from colony B, 3 input libraries) produced 1,110,401,018 paired-end reads which survived quality and adapter trimming using Trimmomatic (version 0.38, parameters used: ILLUMINACLIP: 2:15:10 HEADCROP:10 LEADING:20 TRAILING 20 SLIDINGWINDOW:4:15 MINLEN:30). An average of 98.1% of reads mapped to the *Apis mellifera* HAv3.1 genome using bwa mem (version 0.7.17, default parameters). Duplicates were marked and removed prior to further processing using picard (version 2.10.1, average duplication level 30.2%).

Accessibility of genes was assessed in two ways: 1) counting reads mapping to gene regions, defined as the gene body plus 1 kb upstream (representing the putative promoter of each gene), and 2) by mapping reads to accessibility peaks and then assigning these peaks to genes (see next section for details). This dual approach was used because of limitations to both individual methods. The region-based method uses annotations of genes to define regions likely to influence transcription, but also relies on a predicted promoter region and will miss longer promoters and any long-range enhancers of genes. The peak-based method has the ability to identify differential accessibility in enhancers or other areas not limited to gene annotations, but utilizes a model optimized for ChIP-Seq data rather than ATAC-seq, and also requires post-processing mapping of peaks to genes. By using both region and peak approaches, consistency between approaches lends confidence to any inferences made about gene accessibility.

Peaks were called from deduplicated BAM files using MACS2 (version 2.1.1, command: callpeak, with parameters: --nomodel -g 2.5e8 --nolambda --keep-dup all --slocal 10000) using the appropriate colony and sequencing batch input as control. Peaks were called on each colony and behavioral group separately, then merged and sorted using BEDTools (version 2.26.0, sort and merge commands). This resulted in a total of 11,614 merged peaks with an average width of 721 bp. For both gene region and peak analyses, mapped reads were counted to the respective features using featureCounts from the Subread package (version 1.5.2). An average of 53.3% of

reads were mapped to gene regions, whereas 51.0% of reads were mapped to called peaks as below.

Gene region and peak counts were imported into R for differential accessibility analysis using edgeR. Regions and peaks with less than 1 CPM in at least 2 samples were removed, and remaining count values were normalized using the TMM method. Tagwise dispersion estimates were followed by quasi-likelihood F tests for each pairwise comparison of groups. Differentially accessible region (DAR) and peak (DAP) results for each pairwise comparison are given in Appendix Dataset C.2.

Functional annotation of differential expression and chromatin accessibility

Differential expression

DEG lists were functionally annotated using Gene Ontology (GO) by first mapping putative orthologues between *Apis mellifera* and *Drosophila melanogaster* using reciprocal best BLASTP hits (e-value cutoff = 1e-5). Only DEGs with putative *D. melanogaster* orthologues were included for GO enrichment, and the background list used was all tested genes (those which passed the minimum expression threshold) with putative *D. melanogaster* orthologues. Enrichment tests for biological processes were conducted using GOrilla (Eden et al., 2009) with all significant DEGs (FDR-corrected p-value < 0.05) against the background list, and significantly enriched GO terms (FDR q-value < 0.05) were visualized using wordclouds.com. GO enrichment results for all DEG lists are given in Appendix Dataset C.3.

In addition to GO annotation, DEG lists were compared with multiple previously published studies in social insects. Brain gene expression datasets comparing queens and workers were obtained for *Apis mellifera* (Grozinger et al., 2007), *Megalopta genalis* (Jones et al., 2017), and *Polistes metricus* (Toth et al., 2010). Brain DEG lists comparing reproductive and sterile workers were obtained for *Apis mellifera* (Grozinger et al., 2007) and *Bombus terrestris* (Marshall et al., 2019). Additional brain gene expression datasets for *Apis mellifera* used for comparison were two studies of nurse and forager workers (Alaux et al., 2009a; Whitfield et al., 2003), and a study which measured brain gene expression changes associated with peripheral knockdown of *vitellogenin* via RNAi (Wheeler et al., 2013). Finally, three studies of selection in bees were used to test for enrichment of DEG lists with genes under selection: Woodard et al. (2011) identified genes rapidly evolving in highly eusocial lineages of bees, Harpur et al. (2014) identified genes with signatures of positive selection in *Apis mellifera*, and Kapheim et al. (2015)

identified genes undergoing positive selection across two independent origins of eusociality based on the genomes of 10 bee species.

Putative orthologues between species were identified using BLAST reciprocal best hits (RBH, e-value<10e-5) between predicted peptides (*A. mellifera*, *B. terrestris*) or translated nucleotides (*M. genalis, P. metricus*) against the peptides of *A. mellifera* genome version HAv3.1. Conversion lists between different *A. mellifera* annotation versions and RBH results are provided in Appendix Dataset C.4. For Representation Factor (RF) gene overlap tests, only genes (or putative orthologues) tested in both studies were compared between any two given studies. Significance values were calculated using the hypergeometric function phyper in R. Gene lists and complete RF results are given in Appendix Dataset C.4.

Differential accessibility

Differentially accessible gene regions (DARs) were annotated using GO as for DEG lists above, using all genes which were accessible above the minimum threshold with putative orthologues in *D. melanogaster*. To functionally annotate DAPs, the midpoint coordinate of the 11,614 peaks identified with MACS2 were assigned to genes based on proximity to honey bee gene features (*Apis mellifera* HAv3.1 genome). The following features were considered per gene: promoters (1 kb upstream), introns, exons, 5’ UTR, 3’ UTR, upstream (10 kb upstream), and downstream (10 kb). Peaks not associated with any gene feature were classified as intergenic. When peaks were associated with multiple genes (e.g. the intron of one gene and the promoter of another), they were assigned to individual genes based on the following priority: promoter (highest priority), exon, 5’ UTR, 3’ UTR, intron, upstream, downstream (lowest priority). If a peak was present in the same highest priority class for multiple genes, it was randomly assigned to one gene. In this way, each peak was assigned to either a single gene or considered intergenic. Of the 11,614 peaks, 1822 were assigned to the promoter region of a gene, 776 to exons, 1326 to 5’ UTRs, 273 to 3’ UTRs, 4666 to introns, 1155 to upstream regions, 773 to downstream regions, and 823 peaks were located in intergenic regions.

As before with GO enrichment for DEGs and DARs, DAPs were functionally annotated by mapping peak-associated genes to putative orthologues in *D. melanogaster* using BLASTP. The background list for enrichment analyses was the list of peaks which met the minimum accessibility count threshold for edgeR analysis and which had putative orthologues in *D. melanogaster*. GOrilla (Eden et al., 2009) was used for enrichment tests, and significantly

enriched GO terms were visualized using wordclouds.com. GO enrichment results for all DAR and DAP lists are given in Appendix Dataset C.5.

Results

Specialized behavioral groups are highly transcriptionally distinct

Consistent with stable differences in behavior (Fig. 6.1), foragers (F) and egg-layers (L) exhibited widespread differences in brain gene expression. Nearly half (46%) of all genes expressed in the brain were differentially expressed between F and L groups (FDR corrected $p<0.05$). Generalists (G) shared transcriptional profiles of both F and L, with nearly all genes differentially expressed between G and either specialized group also present on the F vs. L DEG list (Fig. 6.2A). F vs. L DEGs were enriched for cytoplasmic translation and transport gene ontology (GO) biological processes, along with many metabolic and biosynthetic processes (Fig. 6.2B, Appendix Dataset C.3). All but one enriched GO term (114 of 115) were for genes more highly expressed in F relative to L (F-biased genes). The only GO term enriched in L-biased genes relative to F, cytoplasmic translation, was also the only enriched GO term for genes overexpressed in G relative to F. Similarly, GO terms enriched in G-biased genes (relative to L) included many of the transport terms enriched among F-biased genes (Appendix Dataset C.3).

The DEG list for specialized behavioral groups (F vs. L DEGs) was similar to previously published caste-related DEGs in multiple species of social insects (Fig. 6.3). In comparison with two previous studies of honey bee workers, F-biased genes in LW colonies showed significant overlap with F-biased genes in queenright colonies (upregulated in F relative to nurses, N) (F>L, F>N: RF=1.7 p=1.707e-09 Alaux et al., 2009a; RF=1.9 p=1.740e-07 Whitfield et al., 2003). Conversely, L-biased genes overlapped with genes upregulated in queenright N relative to F (RF=1.7, p=3.656e-10 Alaux et al., 2009; RF=2.0, p=1.116e-13, Whitfield et al., 2003). F vs. L DEGs were also enriched for queen (Q) vs. worker (W) brain DEGs in *Apis mellifera* (Grozinger et al., 2007) as well as the facultatively eusocial bee, *Megalopta genalis* (Jones et al., 2017), and the primitively eusocial wasp, *Polistes metricus* (Toth et al., 2010). F-biased genes were enriched for Q-biased genes in *A. mellifera* (RF=1.4, p=5.495e-08), as well as W-biased genes in *M. genalis* (RF=1.6, p=0.010) and *P. metricus* (RF=2.6, p=2.880e-04). L-biased genes were enriched for W-biased genes in *A. mellifera* (RF=1.2, p=0.008) and Q-biased genes in *M. genalis* (RF=2.5, p=0.003). In comparison with studies of reproductive vs. sterile workers (RW vs. SW;

Grozinger et al., 2007; Marshall et al., 2019), L-biased genes were significantly depleted of RW-biased genes in *A. mellfera* (RF=0.6, p=0.037) and enriched for SW-biased genes of *Bombus terrestris* (RF=1.5, p=0.006), while F-biased genes were enriched for RW-biased genes in *B. terrestris* (RF=1.5, p=0.002). F-biased genes were also enriched for genes more highly expressed in control samples of a *vitellogenin* RNAi experiment in *A. mellifera* (RF=1.4, p=5.107e-11; Wheeler et al., 2013).

Additionally, F vs. L DEGs were enriched for genes identified as under selection in two studies of social evolution, suggesting there may be selection for transcriptional divergence related to behavioral specialization (Fig. 6.3). F vs. L DEGs overlapped significantly with genes undergoing positive selection in honey bees (RF=1.1, p=0.015; Harpur et al., 2014) and across highly eusocial species relative to solitary or primitively eusocial species (Woodard et al., 2011). Genes under selection in highly eusocial lineages were enriched specifically for those upregulated in F relative to L (RF=1.4, p=0.009), but not for L-biased DEGs (p=0.106). F vs. L DEGs were not significantly enriched for genes that were identified in a third study as under selection in social lineages of bees (Kapheim et al., 2015), although 30 genes were common to both lists (Appendix Dataset C.4).

Many of the F vs. L DEGs (1480) were differentially expressed or under selection in at least one of the above studies of caste or social-related molecular evolution (Appendix Dataset C.4). Thirty-six genes were implicated in at least four other studies, suggesting strong conservation of these genes in caste-related functions across species with importance for eusocial evolution (Appendix Table B.4). Eighteen of these 36 genes were annotated with functions related to metabolism, and five were members of the glycolysis/gluconeogenesis KEGG pathway, which displayed a strong bias for upregulation in F (Fig. 6.4). Six of the remaining genes not annotated with metabolic functions were related to chromatin processing, including one transcription factor (*cwo*), two ribosomal proteins, and a nucleosome assembly protein.

Brain gene expression is correlated with continual behavioral variation

Pairwise analyses showed that G share transcriptional profiles of both L and F (Fig. 6.2A), raising the possibility that G are intermediate in brain gene expression relative to either specialized group. Principal component analysis (PCA) was used to reduce dimensionality of the transcriptional data, and revealed a nearly continuous gradient of gene expression among individuals. PCs 1 and 2, which explained 31.1% and 11.9% of the total variance in gene

expression, respectively, were significantly correlated with the behavioral specialist score of individual sequenced bees (Fig. 6.5A). G showed intermediate values of these PCs, consistent with an intermediate transcriptional profile. Genes with extreme principal component loading values (upper and lower 5% of loadings) for PC1 were enriched for transmembrane and ion transport, functions related to aerobic and cellular respiration, and energy transport (Fig. 6.5B, left panel). PC2 extreme loading genes were enriched for many processes relating to detection of light, phototransduction, and sensory perception (Fig. 6.5B, right panel). Extreme loadings for both PC1 and PC2 overlapped significantly with DEGs in the pairwise comparison of L and F (PC1: RF=1.2, p=1.393e-06, PC2: RF=1.7, p=8.392e-91).

Brain chromatin accessibility is correlated with patterns of brain gene expression

To test the possibility that the extensive differences in transcriptional profiles between specialized behavioral groups are due to underlying differences in chromatin structure, we used the Assay for Transposase-Accessible Chromatin using sequencing (ATACseq; Buenrostro et al., 2013) to measure accessible chromatin in the brains of the same bees. On a gene by gene basis, open chromatin was significantly correlated with gene expression across all three group of bees, with analyses that included the gene body and putative promoter region of each gene (Fig. 6.6, Pearson's ρ=0.51, p<0.0001). A similar trend was observed when the analyses used accessible peaks, rather than gene regions, although this correlation was not significant (Pearson correlation, p=0.05166). However, genes with nearby chromatin accessibility peaks were, on average, more highly expressed than genes with no assigned peaks (Fig. 6.7, Kolmogorov-Smirnov test, p<0.0001), suggesting peak presence (but not necessarily peak height) weakly influenced expression levels.

Specialized behavioral groups exhibit differences in accessible chromatin

L and F showed differences in accessible chromatin in the brain. 1146 genes had significant (FDR corrected p<0.05) differences in regional accessibility (DARs, Fig. 6.8A). Genes with F-biased DARs (relative to L, 628 genes) were enriched for GO functions related to neuropeptide signaling, metabolic processes, and fat body and adipose tissue development (Appendix Dataset C.5). Genes with L-biased DARs (relative to F, 518 genes) were significantly enriched for just one GO term, regulation of mitochondrial membrane potential (Appendix Dataset C.5). No DARs were found in F relative to G, nor in G relative to L (Fig. 6.8A).

Differences in chromatin accessibility associated with behaviorally related differences in brain gene expression were also apparent when analyzed on the basis of peaks rather than regions. 1794 differentially accessible peaks (DAPs) were identified between F and L, proximal to 1207 genes. No DAPs were identified for F relative to G, and there were only 16 DAPs (assigned to 13 genes) for G relative to L (Fig. 6.8B). These 13 genes also had DAPs for F relative to L (Fig. 6.8B). DAPs between F and L were enriched for 148 GO terms, including developmental processes, morphogenesis, and metabolism (Appendix Dataset C.5). Similar to differentially expressed genes, GO enrichment signal came from those DAPs with a bias in F (i.e. more accessible in F relative to L); no significantly enriched GO terms were identified from L-biased peaks, despite 44% of differential peaks being more accessible in L.

DARs and DAPs were highly concordant, with 2.3-fold more overlap than expected (p=8.942e-67) when considering all genes tested for differential chromatin accessibility. The overlap was also strongly directionally concordant, with 3.9-fold more overlap between F-biased accessibility (p=9.264e-47) and 4.4-fold more overlap in L-biased accessibility (p=2.928e-93) than expected by chance (Fig. 6.8C).

Despite the relatively large numbers of genes displaying differences in chromatin accessibility, accessibility differences between L and F were subtle. Rather than on/off peaks of accessibility, small logFC differences in accessibility were present between behavioral groups (e.g., see Fig. 6.9, Appendix Fig. A.9).

Differential chromatin accessibility is consistent with gene expression differences and correlated with behavioral specialist score

DARs for F relative to L were enriched for F vs. L DEGs. This overlap was directionally-specific: DARs more accessible in F overlapped significantly with genes more highly expressed in F (RF=1.5, p=5.067e-09), and vice versa for L-biased genes (RF=1.3, p=7.444e-04; Fig. 6.10A). Overlapping DARs and DEGs were not enriched for any GO biological processes.

Similar to DARs, DAPs with greater accessibility for F relative to L were enriched near genes more highly expressed in F (RF=1.2, p=0.010; Fig. 6.10B). Unlike the lack of GO enrichment for DAR-overlapping genes, F-biased genes with peak accessibility differences were enriched for multiple GO terms, including axon guidance, neuron projection, and locomotory behavior (Appendix Dataset C.5). L-biased genes had no significant enrichment for DAPs more

accessible in L (p=0.0760), but 168 genes were both more accessible and more highly expressed in L relative to F.

Overall, 862 genes were differentially expressed and had nearby DARs or DAPs between F and L (549 with directional concordance; Appendix Dataset C.6), and 136 genes were differentially expressed and were associated with both DARs and DAPs (85 with directional concordance; Appendix Dataset C.6). Figure 6.9 shows the chromatin accessibility near a gene with directionally concordant accessibility and expression differences between F and L (*Hsp90*, which is more accessible and upregulated in L). Among the genes both differentially expressed and accessible between F and L were 5 genes encoding histones, one histone demethylase, *synaptotagmin-4* and *syntaxin-12* (Appendix Table B.5).

PCA of chromatin accessibility data revealed PCs that were correlated with behavioral variation. PCs for both gene region accessibility (Fig. 6.11A) and peak accessibility (Fig. 6.11B) were identified that were correlated with behavioral specialist score. Gene region accessibility PCs 2 and 4 were correlated with specialist score, and extreme loading genes (upper and lower 5% of loadings) for PC2 were enriched for aspartate family amino acid catabolic process and chitin metabolic process GO terms. PC4 extreme loading genes were not enriched for any GO biological processes.

Peak accessibility PCs 2, 3, and 4 were all significantly correlated with behavioral specialist score (Fig. 6.11B). Genes with extreme PC loading values for each correlated PC showed enrichment for multiple GO terms, including biological processes related to cell-cell adhesion, locomotion, axon guidance, neuron projection guidance, and synapse organization (Appendix Dataset C.5). Many GO terms (11 of 37) were enriched for extreme loading genes of both PCs 2 and 3, while synapse organization was the only enriched term for loadings of PC4.

Discussion

In chapter 5, I described the existence of multiple ancestral-like forms of reproductive plasticity in queenless laying worker (LW) honey bee colonies. Here, I utilized the socio-behavioral plasticity present in these colonies to explore how flexibility in gene regulation may contribute to this variation in social organization. By comparing the brain transcriptomic and accessible chromatin profiles of bees in LW colonies, I discovered continuous molecular variation underlying extensive behavioral plasticity across social phenotypes. In addition, I

identified previously implicated and conserved genes and gene pathways that appear to be co-opted in the regulation of socially-relevant phenotypes across social insect species that are also associated with LW behavioral plasticity within honey bees. These results suggest that the social variation observed in LW colonies may be mediated by similar mechanisms to those important for the evolution and maintenance of eusociality, and may provide glimpses into the types of gene regulatory changes important in transitional stages of social behavior.

Egg-layers and foragers in LW colonies displayed stable differences in the performance of different behaviors. Based on these long-term behavioral differences, I predicted that the brain transcriptional profiles of these groups would be distinct and regulated through stable differences in accessibility chromatin. Consistent with this prediction, I found widespread differences in brain gene expression between specialized layers and foragers, as well as moderate differences in brain chromatin accessibility. Forager-biased genes showed enrichment for a number of metabolic processes, including the glycolysis pathway, which has been previously implicated in caste determination and eusocial origins across bees and wasps (Berens et al., 2014; Jones et al., 2017; Woodard et al., 2011). In addition, despite the queenless social environment, foragers in LW colonies showed transcriptomic signatures similar to foragers from queenright colonies. This result is surprising considering physiological differences between queenright foragers and foragers from LW colonies, including ovary development in nearly half of the foragers from LW colonies sequenced in this study (Appendix Table B.2). Honey bees, like other organisms, are known to exhibit extensive cross-talk between tissues, with changes in peripheral hormone levels leading to behavioral and transcriptomic changes (Robinson, 1985; Schulz et al., 2002; Whitfield et al., 2006). For example, abdominal knockdown via RNAi of the yolk protein precursor and signaling molecule, vitellogenin, leads to considerable transcriptomic changes in the honey bee brain (Wheeler et al., 2013) and precocious behavioral maturation (Nelson et al., 2007). The similar transcriptomic profiles of foragers in LW colonies and those of queenright colonies suggests that despite likely differences in hormone signaling, including any effects of activated ovaries, foraging experience itself may influence brain gene expression, leading to similarities in the brain gene expression profiles of foragers across contexts.

In addition to overlap with queenright forager brain gene expression, the transcriptomic profile of specialized foragers in LW colonies resembled transcriptional signatures of many caste-related genes in other species, as well as genes under selection across lineages of social

insects. This suggests that foraging-related phenotypes are regulated by similar mechanisms across the multiple independent origins of the worker caste.

Intriguingly, egg-layers resembled queenright nurse honey bees with respect to brain transcription. LW were previously found to have enlarged hypopharyngeal glands (Naeger et al., 2013), which are developed in nurse bees and produce glandular secretions to feed larvae, consistent with the transcriptomic similarities presented here. However, layers in my colonies had very little, if any, ability to engage in nursing behavior. When bees were collected from the LW colonies for sequencing, only one colony of six had begun to rear drones, and only a small number of larvae were present (see Chapter 5). This suggests that mechanistically, LW may be primed to engage in nursing behavior, and the gene expression profile of specialized layers reflects this priming. Evolutionarily, layers may resemble nurse bees in brain gene expression because similar gene networks are involved in both nursing and reproductive behaviors. This latter suggestion is consistent with proposed links between reproductive signaling and brood care in honey bees, an extension of the ovarian ground plan hypothesis (OGPH) (Amdam et al., 2006b; Graham et al., 2011; Page et al., 2012; Wang et al., 2011). While the OGPH suggests a decoupling of reproductive and non-reproductive behaviors in the evolution of queen and worker castes from a solitary ancestor (West-Eberhard, 1987; West-Eberhard, 1996), this conceptual extension suggests that maternal and reproduction-related signaling pathways were co-opted for worker subcaste specializations (Amdam et al., 2006a; Page et al., 2006). Links between reproductive signaling and non-reproductive behaviors, including maternal care, have been found in other species of bees and wasps (Kapheim and Johnson, 2017; Toth et al., 2007), supporting the idea that during the multiple evolutionary origins of social behavior, regulation of non-reproductive traits was co-opted from ancestral reproductive signaling pathways. Together, these results suggest that LW reproduction may be regulated by ancestral gene regulatory networks and representative of a reversion to a behavioral state that occurred earlier in social evolution.

Many genes with both expression and chromatin accessibility differences between specialized layers and foragers were annotated with functions in metabolic pathways. In a comparative study across lineages of bees with variation in social traits, genes in the glycolysis pathway were identified as rapidly evolving in advanced eusocial lineages (Woodard et al., 2011). Differences in the expression of glycolysis enzymes also appear to be associated with LW phenotypes, suggesting that carbohydrate metabolism may be a conserved regulator of social

plasticity in bees across contexts. Within a halictid bee with a genetically determined social polymorphism, metabolic genes in the TOR signaling and insulin-like growth factor receptor signaling pathways were found to contain either coding changes or nearby regulatory variants between social forms (Kocher et al., 2018). This provides further support for a general role of metabolic processes in regulating social plasticity, across lineages (Woodard et al., 2011), intraspecific populations (Kocher et al., 2018), and even within colonies of social insects (this study).

In addition to genes related to metabolism, a number of genes encoding histone proteins were both differentially accessible and expressed between specialized groups. Histone modifications have been implicated in queen-worker caste determination (Wojciechowski et al., 2018), as well as in modulating responses to socially relevant stimuli within the worker caste of honey bees (Shpigler et al., 2017a; Shpigler et al., 2018). A number of additional chromatin processing genes were differentially expressed in LW behavioral groups, including genes involved in caste- or subcaste-related phenotypes in other contexts, or with signatures of positive selection associated with social behavior in bees (Harpur et al., 2014; Kapheim et al., 2015; Woodard et al., 2011). These results suggest that chromatin-associated mechanisms of phenotypic plasticity play a prominent role in establishing a variety of phenotypes in social insects and other species (Snyder-Mackler et al., 2018).

Forager-biased genes were enriched for GO functions related to neuropeptide signaling, axon guidance, and neuron projection. Additionally, genes with chromatin accessibility variation which contributed to behavioral variation based on PC analyses were enriched for neuron projection guidance and synapse organization. These results suggest a role of neuronal reorganization and plasticity in mediating LW phenotypes. Foragers showed upregulation of both *synaptotagmin-4* and *syntaxin-12,* as well as differential accessibility of these genes relative to specialized egg-layers. *syntaxin* genes encode products which are components of the SNARE complex of proteins which are involved in membrane fusion events (Bennett et al., 1993; Chen and Scheller, 2001). SNARE proteins are best known for their role in docking of synaptic vesicles in neurons (Hanson et al., 1997), where synaptotagmin proteins may act as calcium sensors leading to exocytosis and neurotransmitter release (Chapman, 2008; c.f. O'Connor and Lee, 2002). s*yntaxin 1a* has been associated with variation in social behavior in multiple species, including migratory locusts (solitary vs. gregarious forms, Chen et al., 2015), socially

polymorphic halictid bees (solitary vs. social populations, Kocher et al., 2018), and unusual social deficits in *syx1a*-knockout mice (Fujiwara et al., 2016). While the specific roles of *syntaxin-12* and *synaptotagmin-4* are unknown in honey bees, these findings raise the possibility that variation in LW social behavior is mediated through highly conserved mechanisms of neuronal plasticity.

In addition to widespread differences in gene expression and chromatin accessibility between specialized behavioral groups, the discovery of generalist bees that performed both reproductive and non-reproductive behaviors (Chapter 5) allowed me to explore in this chapter what kind of molecular mechanisms might subserve this ancestral-like flexibility in behavior. To enable performance of both reproductive and non-reproductive behaviors, one prediction was that generalists would display unique brain gene expression and chromatin accessibility profiles compared with bees performing either reproductive or non-reproductive behaviors but not both. Instead, I found that both brain gene expression and chromatin profiling of generalist bees was intermediate between specialized layers and foragers, with very few transcripts uniquely expressed in generalist bees, and almost no significant differences in chromatin accessibility between generalists and either specialized group. These results suggest that at the molecular level, a reproductive division of labor may readily emerge from a solitary ancestral network simply by tweaking the timing and degree of expression of key genes in two directions, one toward reproductive specialization and one toward non-reproductive specialization (Linksvayer and Wade, 2005; West-Eberhard, 1987; West-Eberhard, 1996). The accessible chromatin landscape of the genes in this putative solitary ancestral network, much like in the generalist bees, would allow for performance of multiple behaviors. Variation in this ancestral profile may have been selected upon, leading to a bifurcation of accessibility and downstream gene expression toward specialized castes through disruptive selection (Eberhard, 1980; West-Eberhard, 1987).

At first glance, the large number of differentially expressed genes between nurses and foragers in queenright colonies (Alaux et al., 2009a; Whitfield et al., 2003) or between the layers and foragers presented here would indicate highly discretized behavioral states of the honey bee worker. However, the brain gene expression profile of generalist bees suggests a different perspective. Generalists display intermediate behavior as well as intermediate transcriptional and chromatin landscapes, which suggests that seemingly discretized states are subserved by gene

expression profiles and gene regulatory mechanisms that are continuous and not discretized. This continuous nature of both molecular and behavioral profiles suggests a high degree of fluidity in honey bee phenotypes, despite their long evolutionary divergence from a solitary ancestor (~75 my, Branstetter et al., 2017). In another highly eusocial species, the carpenter ant *Camponotus floridanus,* surprising plasticity in the behavior of morphologically specialized subcastes of workers was demonstrated (as in Wilson, 1980), with manipulations of histone modifications able to induce shifts in behavior in one subcaste that is typically associated with another subcaste (Simola et al., 2015). Together, these results suggest that even in highly derived social systems, a broad range of phenotypes may be latent and available given the appropriate environmental or social cues.

Compared with chromatin changes related to development, where widespread peaks of accessibility are present or absent in different cell types (Ackermann et al., 2016) or across developmental stages (Lu et al., 2016), differences in chromatin accessibility observed between honey bee worker behavioral groups were subtle. This may reflect a biological difference between developmental and behavioral plasticity, with stable chromatin in a given tissue type and life stage laying the groundwork for transcriptional and neural plasticity to drive behavioral differences. Consistent with this idea, a previous study of honey bee behavior using ChIP-seq revealed a small number (<200) of differential peaks for H3K27ac enrichment (a histone mark of open chromatin and active regulatory elements) following exposure to a socially relevant stimulus (Shpigler et al., 2018). In contrast, differences in histone modifications between queen and worker castes during larval development were more substantial (Wojciechowski et al., 2018), although still not at the level of cell type-specific chromatin marks that have been observed in other studies (Ernst et al., 2011; Xi et al., 2007).

The subtle differences observed in honey bee studies may also be a result of technical limitations when using homogenate tissues (whole brains or brain pools in this study and Shpigler et al., 2018, respectively; whole larval heads in Wojciechowski et al., 2018), or sampling at time points before or after more dramatic chromatin changes. However, the effects of social status on chromatin accessibility in immune cells of female rhesus macaques were also modest in comparison with gene expression changes (Snyder-Mackler et al., 2018), suggesting the difference between behavior and development in degree of chromatin variation is not purely technical, or unique to social insects. Advances in single-cell sequencing will provide insights

into these types of questions, as will studies of chromatin accessibility in additional behavioral contexts.

Overall, the molecular profiling of behavioral plasticity presented here suggests that changes in both accessible chromatin and gene expression enable the surprising behavioral continuum between reproductive and non-reproductive individuals in queenless LW honey bee colonies. Despite the absence of a queen and unique colony social dynamics of these bees, transcriptomic profiles of specialized workers mirror those of queenright honey bee subcastes, consistent with hypotheses suggesting ancestral reproductive signaling pathways were co-opted during the evolution of the worker caste. Further, I find support for a behavioral continuum among bees engaged in a reproductive division of labor, suggesting that honey bee subcastes may have emerged from incremental changes in gene regulation from an ancestral solitary state. Finally, the extent of gene expression variation relative to chromatin changes suggests that behavioral plasticity may be facilitated through extensive transcriptional and neuronal plasticity, with potentially less dependence on longer-term changes in chromatin structural organization. This may reflect a fundamental difference between developmental and behavioral plasticity, with implications for how these different types of phenotypic plasticity may influence and be shaped by evolutionary processes.

Acknowledgements

I would like to thank Amy Cash Ahmed and Vikyath Rao for their contributions as collaborators on this chapter of work.

Figures

Figure 6.1

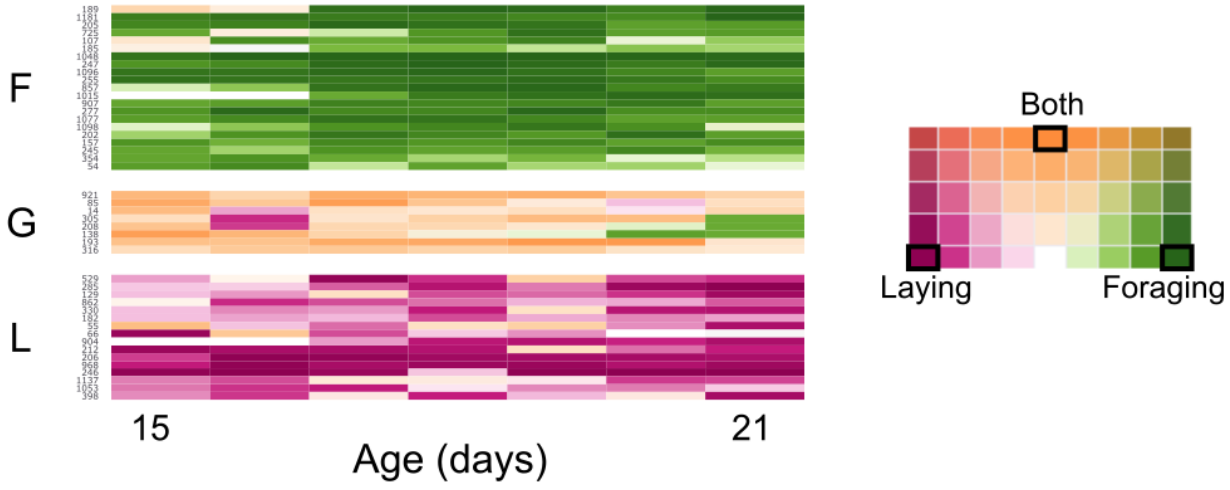


Figure 6.1. Daily specialist and generalist scores represented in 2D color space for all sequenced individuals. Individual bees are represented as rows, and columns represent the age of the colony (and bees). “Both” refers to generalist behavior, with performance of both egg-laying and foraging. F: foragers, L: layers, G: generalists.

Figure 6.2

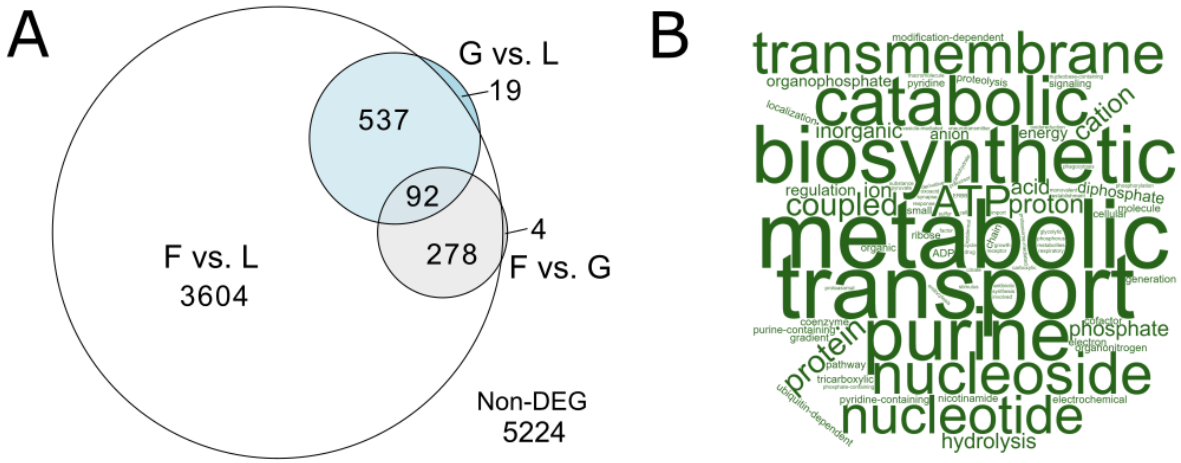


Figure 6.2. (A) Euler diagram for overlaps of pairwise differentially expressed genes (DEGs) between behavioral groups. Note that one gene was overlapping between F vs. G and G vs. L but is not represented in the diagram due to graphical constraints. (B) Word cloud of significantly enriched GO term descriptions for forager-biased differentially expressed genes. The size of the word corresponds with the frequency with which that term appears on the list of significantly enriched GO biological processes. The word “process” was removed from descriptions prior to generation of the word frequency list. F: foragers, L: layers, G: generalists.

Figure 6.3

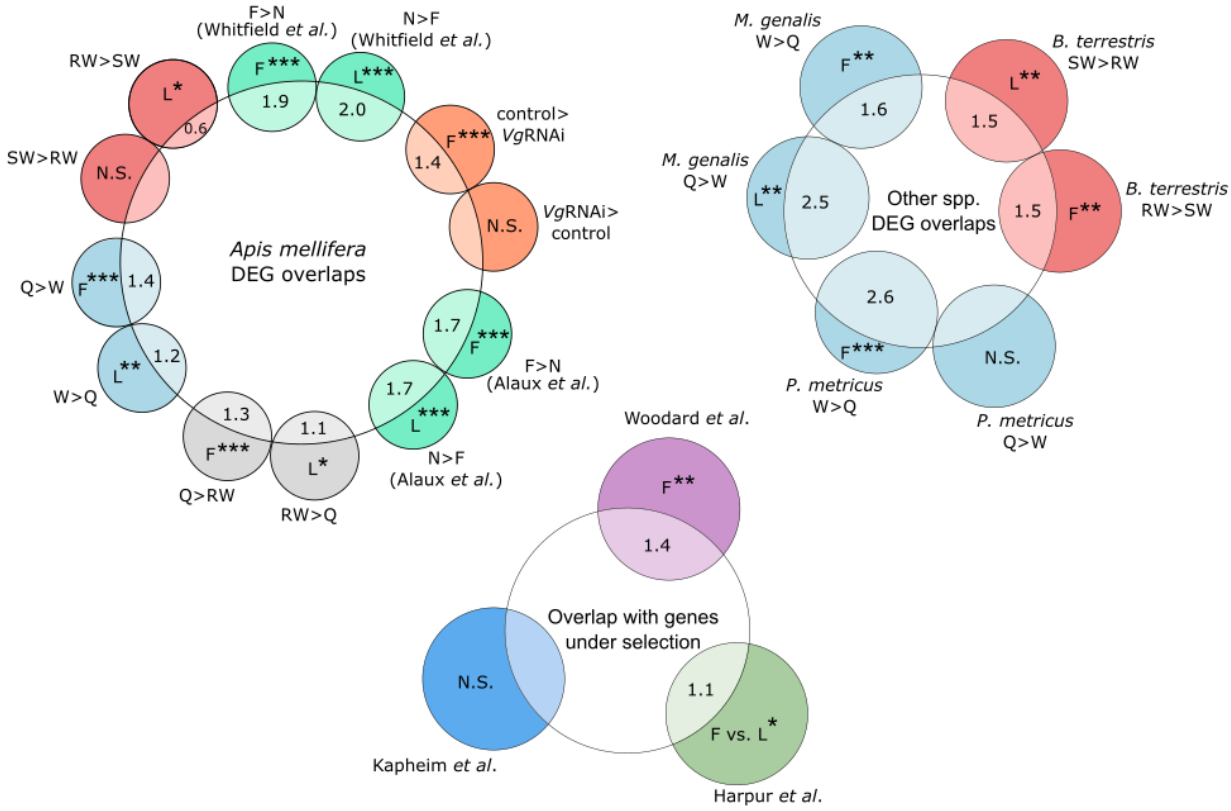
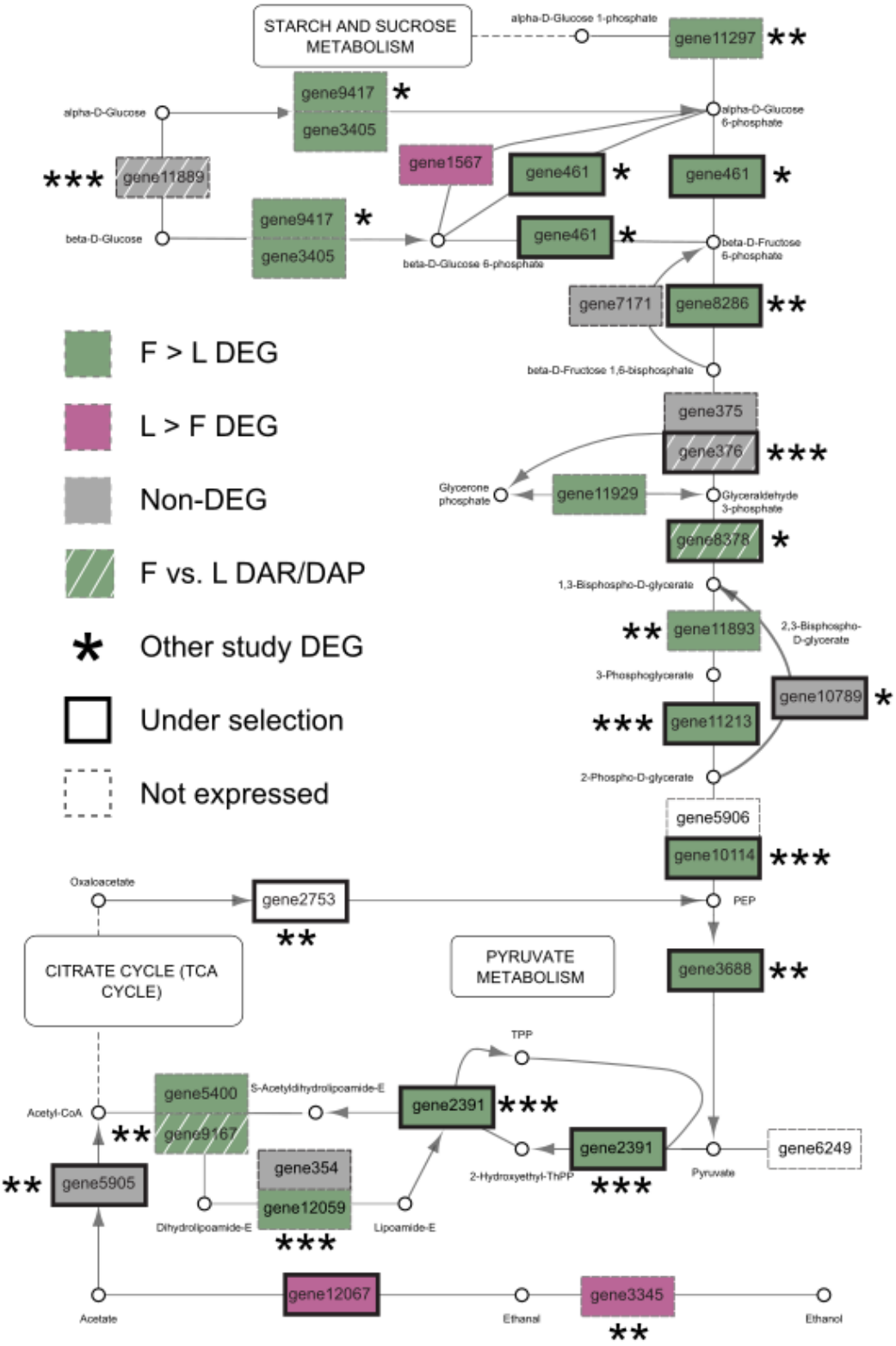


Figure 6.3. Overlaps of forager (F) vs. layer (L) differentially expressed genes (DEGs) with previously published gene expression and molecular evolution datasets (references provided in main text). Representation factors (RF) for each significant comparison are given inside the central circles. An RF>1 indicates more overlap than expected by chance, while an RF<1 indicates less overlap than expected. Significance values based on hypergeometric tests of overlap are given by asterisks: *p<0.05, **p<0.01, ***p<0.001. Letter(s) before asterisks (F, L, or F vs. L) denote whether forager-biased (F), layer-biased (L) or all DEGs (F vs. L) overlapped significantly with the gene list of interest. Colors indicate phenotypes being compared. N: nurse, Q: queen, W: worker, RW: reproductive worker, SW: sterile worker.

Figure 6.4


STARCH AND SUCROSE METABOLISM
alpha-D-Glucose 1-phosphate
gene11297 **
gene9417 *
gene3405
alpha-D-Glucose
alpha-D-Glucose 6-phosphate
*** gene11889
gene1567
gene461 *
gene461 *
gene9417 *
gene3405
beta-D-Glucose
gene461 *
beta-D-Glucose 6-phosphate
beta-D-Fructose 6-phosphate
gene7171
gene8286 **
beta-D-Fructose 1,6-bisphosphate
gene375
gene376 ***
Glycerone phosphate
gene11929
Glyceraldehyde 3-phosphate
gene8378 *
1,3-Bisphospho-D-glycerate
** gene11893
2,3-Bisphospho-D-glycerate
3-Phosphoglycerate
gene10789 *
*** gene11213
2-Phospho-D-glycerate
gene5906
gene10114 ***
F > L DEG
L > F DEG
Non-DEG
F vs. L DAR/DAP
* Other study DEG
Under selection
Not expressed
Oxaloacetate
gene2753 **
PEP
CITRATE CYCLE (TCA CYCLE)
PYRUVATE METABOLISM
gene3688 **
TPP
gene5400
gene9167
S-Acetyldihydrolipoamide-E
gene2391 ***
Acetyl-CoA
** gene5905
gene354
gene12059 ***
Dihydrolipoamide-E
Lipoamide-E
gene2391 ***
2-Hydroxyethyl-ThPP
Pyruvate
gene6249
gene12067
Acetate
Ethanal
gene3345 **
Ethanol

(Figure 6.4 caption continued)

Figure 6.4. Annotated honey bee genes in the glycolysis/gluconeogenesis KEGG pathway (Kyoto Encyclopedia of Genes and Genomes, http://www.genome.jp/kegg/). Genes are colored to indicate whether they are differentially expressed (DEG) between foragers (F) and layers (L), and hashed if they have either a nearby differentially accessible region or differentially accessible peak (DAR/DAP). Numbers of asterisks indicate how many, if any, comparative studies of gene expression in social insect castes or caste-related phenotypes also found the gene differentially expressed (studies discussed in main text and in Fig. 6.3). Thick borders indicate genes which were found to be under selection in at least one of three studies of selection in bees (Harpur et al., 2014; Kapheim et al., 2015; Woodard et al., 2011). “Not expressed” genes refer to those not expressed above the minimum threshold in the current F vs. L dataset, while non-DEG genes were expressed but not differentially expressed in F vs. L at an FDR-corrected $p<0.05$.

Figure 6.5

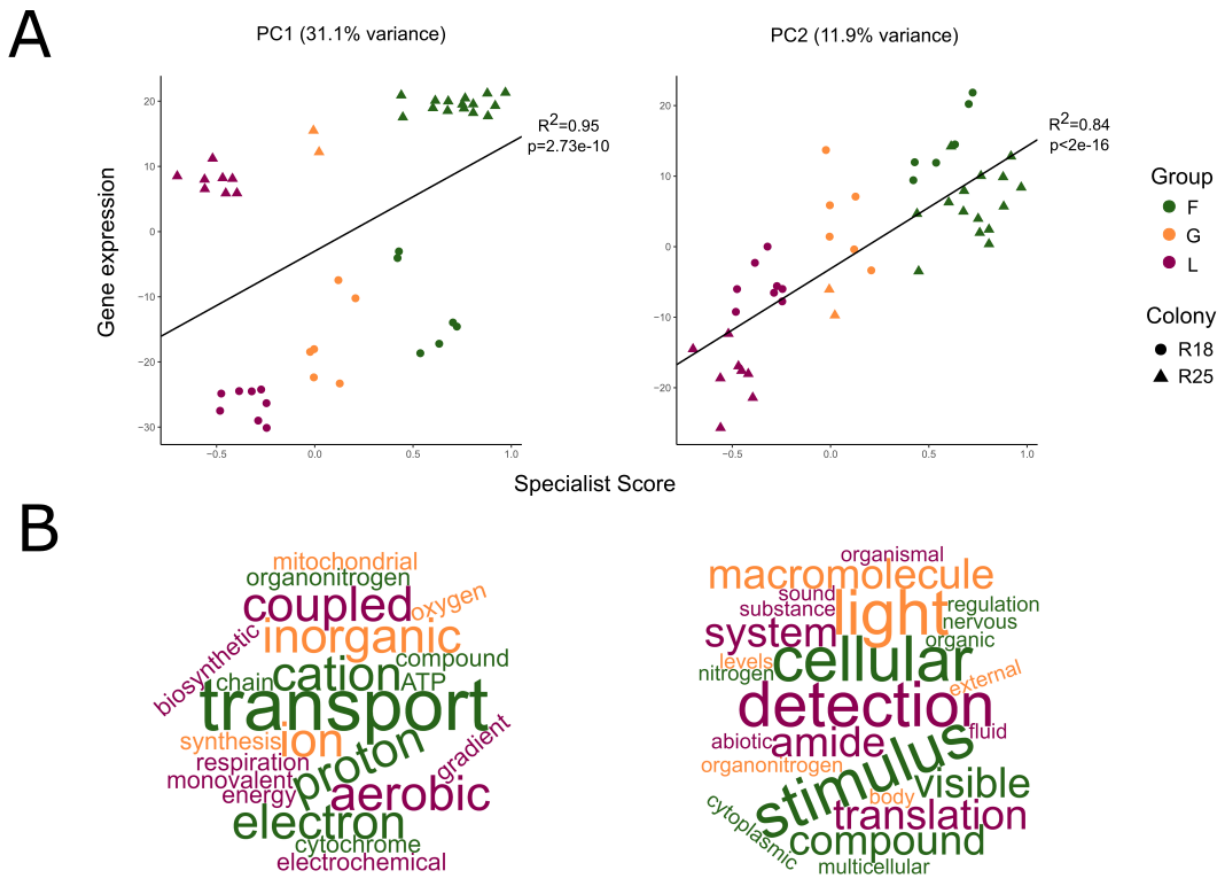


Figure 6.5. (A) Principal components of gene expression are correlated with behavioral specialist score. Correlation coefficients and p-values are given for model which includes specialist score and colony. (B) Word clouds of significantly enriched GO term descriptions for PC1 (left) and PC2 (right) extreme loading genes. The size of the word corresponds with the frequency with which that term appears on the list of significantly enriched GO biological processes. The word "process" was removed from descriptions prior to generation of word frequency lists. F: foragers, G: generalists, L: layers.

Figure 6.6

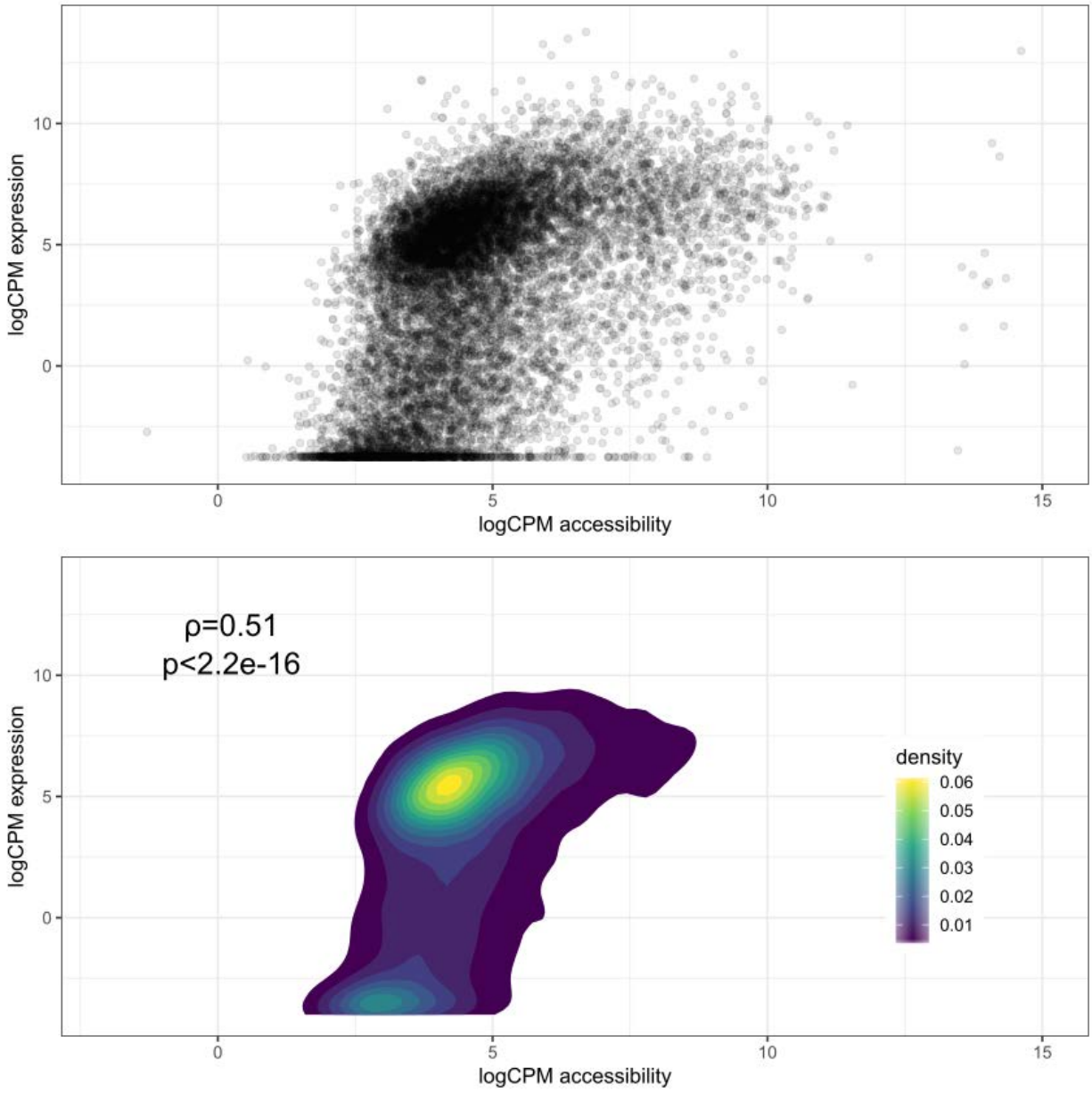


Figure 6.6. Average brain gene expression (logCPM = log counts per million) across samples is correlated with average chromatin accessibility of the gene region (1kb upstream of and including gene body). Upper panel shows scatterplot of individual genes, while lower panel shows relative density of points. Spearman rank correlation test, Spearman's ρ=0.51, p<2.2e-16.

Figure 6.7

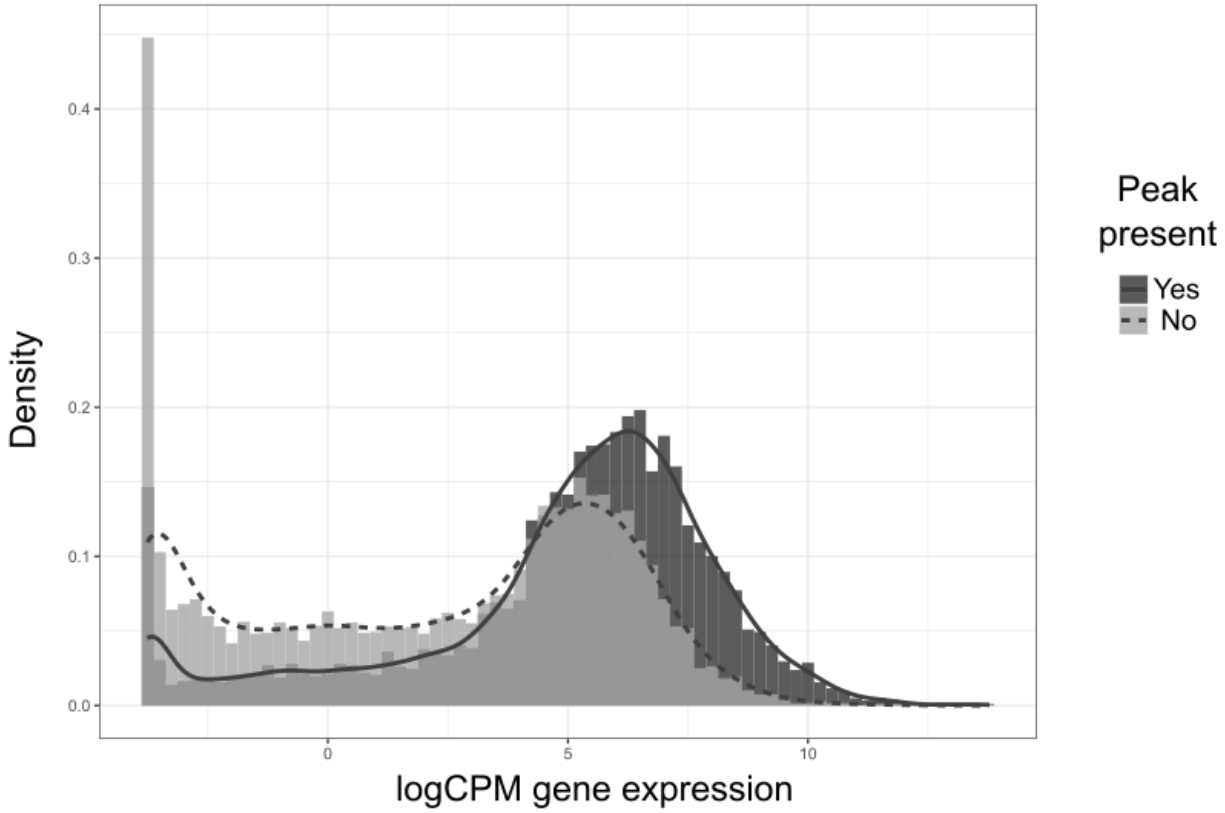


Figure 6.7. Histogram (bars) and density (lines) of gene expression for genes with (light gray) and without (dark gray) nearby peaks of chromatin accessibility. Distributions are significantly different ($p<0.0001$, Kolmogorov-Smirnov test).

Figure 6.8

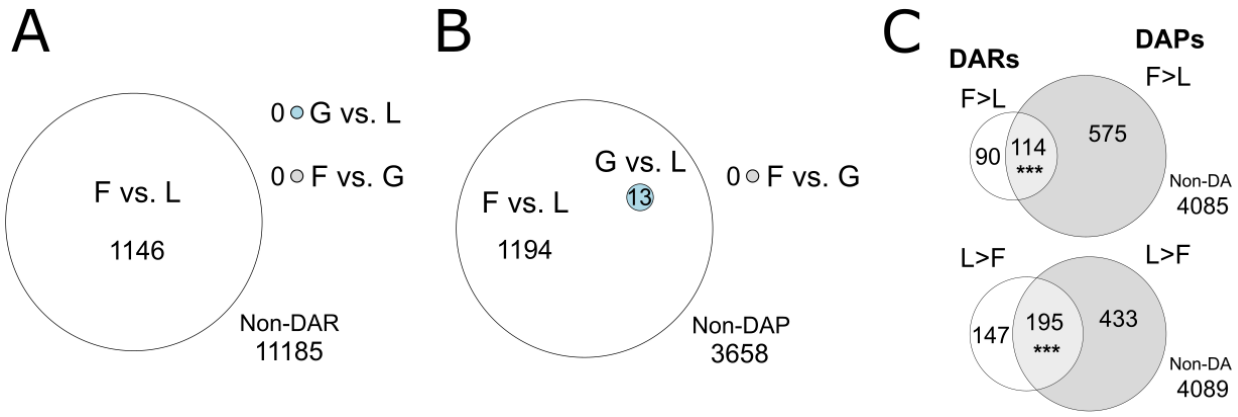


Figure 6.8. Euler diagram for overlaps of pairwise differentially accessible chromatin regions (DARs, A) and differentially accessible peaks (DAPs, B) between behavioral groups. (C) Venn diagrams for overlaps between directional DARs and DAPs between specialized layers and foragers. *** indicates p<0.0001 for hypergeometric tests of overlap. F: foragers, L: layers, G: generalists.

Figure 6.9

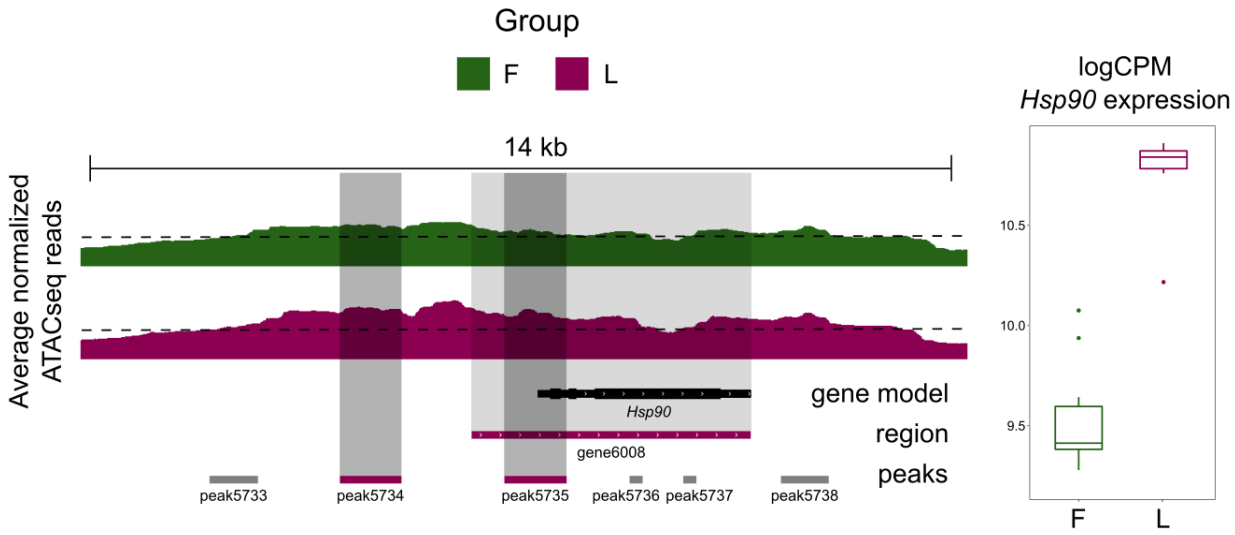


Figure 6.9. Brain chromatin accessibility near *Hsp90*, a gene with differential accessibility (DAR: FDR-corrected p=0.000235, DAPs: FDR-corrected p=0.001854 for peak5734, p=0.000465 for peak5735) and differential expression (DEG: FDR-corrected p=5.68e-12) between specialized foragers (F, green) and layers (L, purple) in colony R25 (colony R18 shown in Appendix Fig. A.9). Color of the gene region and peaks indicates direction of differential accessibility; grey peaks are not differentially accessible. Shaded rectangles highlight differentially accessible peaks or regions. N=23.

Figure 6.10

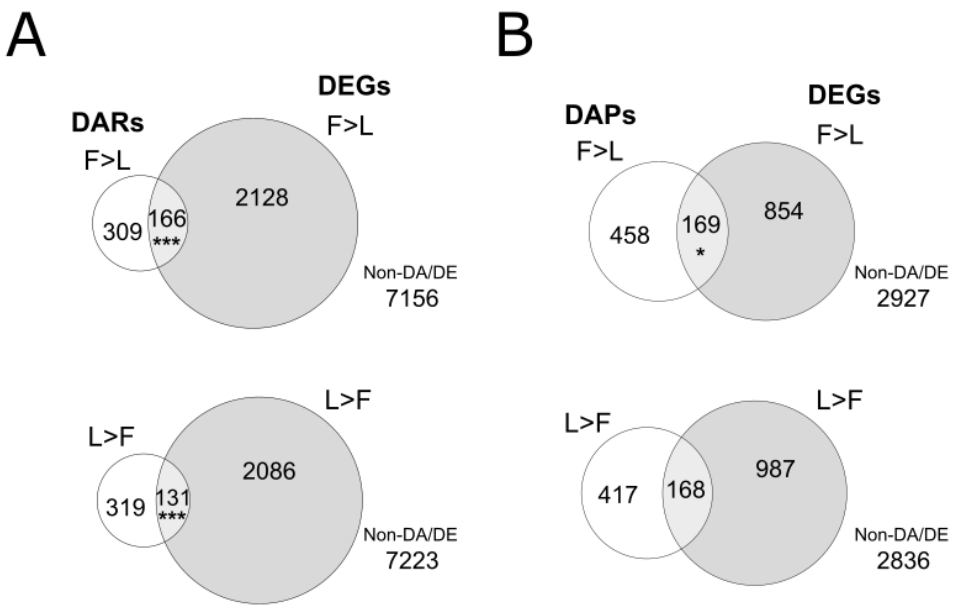


Figure 6.10. Venn diagrams for overlaps of pairwise forager (F) vs. layer (L) differentially accessible chromatin regions (DARs, A) and differentially accessible chromatin peaks (DAPs, B) with differentially expressed genes (DEGs). *p<0.05, *** p<0.0001 for hypergeometric tests of overlap.

Figure 6.11

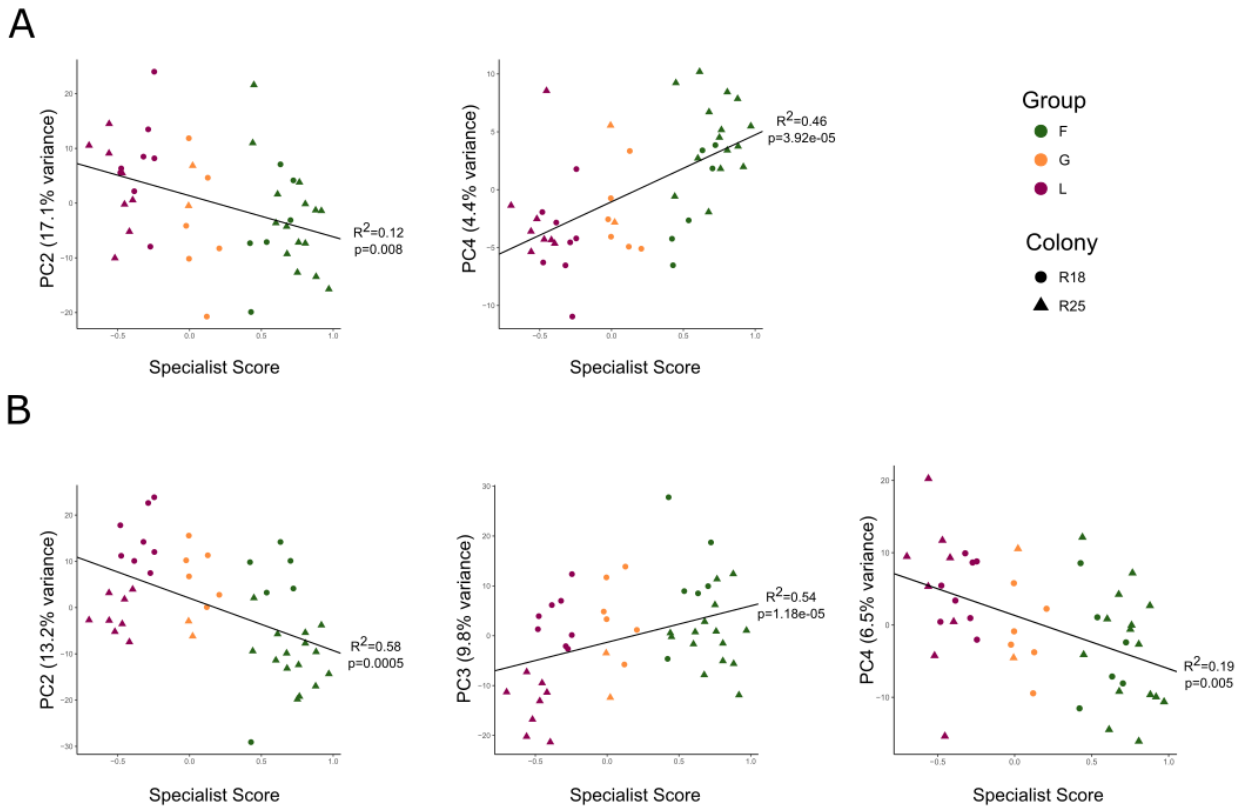


Figure 6.11. (A) Gene region accessibility and (B) chromatin peak accessibility principal components correlated with behavioral specialist score. Correlation coefficients and p-values are given for main effect of specialist score. F: foragers, G: generalists, L: layers.

References

Abzhanov, A., Kuo, W. P., Hartmann, C., Grant, B. R., Grant, P. R. and Tabin, C. J. (2006). The calmodulin pathway and evolution of elongated beak morphology in Darwin's finches. *Nature* 442, 563-567.

Ackermann, A. M., Wang, Z., Schug, J., Naji, A. and Kaestner, K. H. (2016). Integration of ATAC-seq and RNA-seq identifies human alpha cell and beta cell signature genes. *Mol. Metab.* 5, 233–244.

Alaux, C., Le Conte, Y., Adams, H. A., Rodriguez-Zas, S., Grozinger, C. M., Sinha, S. and Robinson, G. E. (2009a). Regulation of brain gene expression in honey bees by brood pheromone. *Genes, Brain Behav.* 8, 309–319.

Alaux, C., Sinha, S., Hasadsri, L. and Hunt, G. J. (2009b). Honey bee aggression supports a link between gene regulation and behavioral evolution. *Proc. Natl. Acad. Sci.* 106, 15400–15405.

Amdam, G. V, Csondes, A., Fondrk, M. K. and Page Jr, R. E. (2006). Complex social behaviour derived from maternal reproductive traits. *Nature* 439, 76-78.

Ament, S., Wang, Y., Chen, C.-C., Blatti, C., Hong, F., Liang, Z., Negre, N., White, K., Rodriquez-Zas, S., Mizzen, C., et al. (2012). The transcription factor *Ultraspiracle* influences honey bee social behavior and behavior-related gene expression. *PLoS Genet.* 8, e1002596.

Ashby, R., Forêt, S., Searle, I. and Maleszka, R. (2016). MicroRNAs in honey bee caste determination. *Sci. Rep.* 6, 18794.

Bannister, A. J. and Kouzarides, T. (2011). Regulation of chromatin by histone modifications. *Cell Res.* 21, 381-395.

Bennett, M. K., Garcia-Arrarás, J. E., Elferink, L. A., Peterson, K., Fleming, A. M., Hazuka, C. D. and Scheller, R. H. (1993). The syntaxin family of vesicular transport receptors. *Cell* 74, 863–873.

Berens, A. J., Hunt, J. H. and Toth, A. L. (2014). Comparative transcriptomics of convergent evolution: Different genes but conserved pathways underlie caste phenotypes across lineages of eusocial insects. *Mol. Biol. Evol.* 32, 690–703.

Bourke, A. F. G. (2011). *Principles of Social Evolution*. Oxford: Oxford University Press.

Branstetter, M. G., Danforth, B. N., Pitts, J. P., Faircloth, B. C., Ward, P. S., Buffington, M. L., Gates, M. W., Kula, R. R. and Brady, S. G. (2017). Phylogenomic insights into the evolution of stinging wasps and the origins of ants and bees. *Curr. Biol.* 27, 1019–1025.

Buenrostro, J. D., Giresi, P. G., Zaba, L. C., Chang, H. Y. and Greenleaf, W. J. (2013). Transposition of native chromatin for fast and sensitive epigenomic profiling of open chromatin, DNA-binding proteins and nucleosome position. *Nat. Methods* 10, 1213–1218.

Bukhari, S. A., Saul, M. C., Seward, C. H., Zhang, H., Bensky, M., James, N., Zhao, S. D., Chandrasekaran, S., Stubbs, L. and Bell, A. M. (2017). Temporal dynamics of neurogenomic plasticity in response to social interactions in male threespined sticklebacks. *PLoS Genet.* 13, e1006840.

Carroll, S. B. (2008). Evo-Devo and an expanding evolutionary synthesis: A genetic theory of morphological evolution. *Cell* 134, 25–36.

Cedar, H. (1988). DNA methylation and gene activity. *Cell* 53, 3–4.

Chandrasekaran, S., Ament, S. A., Eddy, J. A., Rodriguez-Zas, S. L., Schatz, B. R., Price, N. D. and Robinson, G. E. (2011). Behavior-specific changes in transcriptional modules lead to distinct and predictable neurogenomic states. *Proc. Natl. Acad. Sci.* 108, 18020.

Chapman, E. R. (2008). How does synaptotagmin trigger neurotransmitter release? *Annu. Rev. Biochem.* 77, 615–641.

Chen, Y. A. and Scheller, R. H. (2001). SNARE-mediated membrane fusion. *Nat. Rev. Mol. cell Biol.* 2, 98-106.

Chen, Q., He, J., Ma, C., Yu, D. and Kang, L. (2015). Syntaxin 1A modulates the sexual maturity rate and progeny egg size related to phase changes in locusts. *Insect Biochem. Mol. Biol.* 56, 1–8.

Cruz, Y. P. (1981). A sterile defender morph in a polyembryonic hymenopterous parasite. *Nature* 294, 446–447.

Dall, S. R. X., Giraldeau, L. A., Olsson, O., McNamara, J. M. and Stephens, D. W. (2005). Information and its use by animals in evolutionary ecology. *Trends Ecol. Evol.* 20, 187–193.

Eberhard, W. G. (1980). Horned beetles. *Sci. Am.* 242, 166–183.

Eden, E., Navon, R., Steinfeld, I., Lipson, D. and Yakhini, Z. (2009). GOrilla: A tool for discovery and visualization of enriched GO terms in ranked gene lists. *BMC Bioinformatics* 10, 48.

Elango, N., Hunt, B. G., Goodisman, M. A. D. and Yi, S. V. (2009). DNA methylation is widespread and associated with differential gene expression in castes of the honeybee, *Apis mellifera*. *Proc. Natl. Acad. Sci.* 106, 11206–11211.

Emlen, D. J. (2000). Integrating development with evolution: A case study with beetle horns: Results from studies of the mechanisms of horn development shed new light on our understanding of beetle horn evolution. *Bioscience* 50, 403–418.

Emlen, D. J. and Nijhout, H. F. (2000). The development and evolution of exaggerated morphologies in insects. *Annu. Rev. Entomol.* 45, 661–708.

Ernst, J., Kheradpour, P., Mikkelsen, T. S., Shoresh, N., Ward, L. D., Epstein, C. B., Zhang, X., Wang, L., Issner, R., Coyne, M., et al. (2011). Systematic analysis of chromatin state dynamics in nine human cell types. *Nature* 473, 43–49.

Feldmeyer, B., Elsner, D. and Foitzik, S. (2014). Gene expression patterns associated with caste and reproductive status in ants: Worker-specific genes are more derived than queen-specific ones. *Mol. Ecol.* 23, 151–161.

Filipowicz, W., Bhattacharyya, S. N. and Sonenberg, N. (2008). Mechanisms of post-transcriptional regulation by microRNAs: Are the answers in sight? *Nat. Rev. Genet.* 9, 102-114.

Foret, S., Kucharski, R., Pellegrini, M., Feng, S., Jacobsen, S. E., Robinson, G. E. and Maleszka, R. (2012). DNA methylation dynamics, metabolic fluxes, gene splicing, and alternative phenotypes in honey bees. *Proc. Natl. Acad. Sci.* 109, 4968–4973.

Fujiwara, T., Sanada, M., Kofuji, T. and Akagawa, K. (2016). Unusual social behavior in HPC-1/syntaxin1A knockout mice is caused by disruption of the oxytocinergic neural system. *J. Neurochem.* 138, 117–123.

Gibbs, J., Brady, S. G., Kanda, K. and Danforth, B. N. (2012). Phylogeny of halictine bees supports a shared origin of eusociality for *Halictus* and *Lasioglossum* (Apoidea: Anthophila: Halictidae). *Mol. Phylogenet. Evol.* 65, 926-939.

Graham, A. M., Munday, M. D., Kaftanoglu, O., Page, R. E., Amdam, G. V. and Rueppell, O. (2011). Support for the reproductive ground plan hypothesis of social evolution and major QTL for ovary traits of Africanized worker honey bees (*Apis mellifera* L.). *BMC Evol. Biol.* 11, 95.

Grozinger, C. M., Whitfield, C. W., Robinson, G. E. and Sharabash, N. M. (2003). Pheromone-mediated gene expression in the honey bee brain. *Proc. Natl. Acad. Sci.* 100, 14519–14525.

Grozinger, C. M., Fan, Y., Hoover, S. E. R. and Winston, M. L. (2007). Genome-wide analysis reveals differences in brain gene expression patterns associated with caste and reproductive status in honey bees (*Apis mellifera*). *Mol. Ecol.* 16, 4837–4848.

Grunstein, M. (1997). Histone acetylation in chromatin structure and transcription. *Nature* 389, 349-352.

Guhaniyogi, J. and Brewer, G. (2001). Regulation of mRNA stability in mammalian cells. *Gene* 265, 11–23.

Hanson, P. I., Heuser, J. E. and Jahn, R. (1997). Neurotransmitter release - four years of SNARE complexes. *Curr. Opin. Neurobiol.* 7, 310–315.

Harpur, B. A., Kent, C. F., Molodtsova, D., Lebon, J. M. D., Alqarni, A. S., Owayss, A. A., and Zayed, A. (2014). Population genomics of the honey bee reveals strong signatures of positive selection on worker traits. *Proc. Natl. Acad. Sci.* 111, 2614–2619.

Harrison, M. C., Hammond, R. L. and Mallon, E. B. (2015). Reproductive workers show queenlike gene expression in an intermediately eusocial insect, the buff-tailed bumble bee *Bombus terrestris*. *Mol. Ecol.* 24, 3043–3063.

He, X. J., Jiang, W. J., Zhou, M., Barron, A. B. and Zeng, Z. J. (2017). A comparison of honeybee (*Apis mellifera*) queen, worker and drone larvae by RNA-Seq. *Insect Sci.* 00, 1–11.

Herb, B. R., Wolschin, F., Hansen, K. D., Aryee, M. J., Langmead, B., Irizarry, R., Amdam, G. V and Feinberg, A. P. (2012). Reversible switching between epigenetic states in honeybee behavioral subcastes. *Nat. Neurosci.* 15, 1371–1373.

Herb, B. R., Shook, M. S., Fields, C. J. and Robinson, G. E. (2018). Defense against territorial intrusion is associated with DNA methylation changes in the honey bee brain. *BMC Genomics* 19, 216.

Hess, G. (1942). Über den einfluss der weisellosigkeit und des fruchtbarkeitsvitamins e auf die ovarien der bienenarbeiterin. *Schweiz. Bienen Ztg.* 2, 33–110.

Hines, H. M., Hunt, J. H., O'Connor, T. K., Gillespie, J. J. and Cameron, S. A. (2007). Multigene phylogeny reveals eusociality evolved twice in vespid wasps. *Proc. Natl. Acad. Sci.* 104, 3295–3299.

Hobert, O. (2008). Gene regulation by transcription factors and microRNAs. *Science* 319, 1785–1786.

Hölldobler, B. and Wilson, E. O. (2008). *The Superorganism: The Beauty, Elegance, and Strangeness of Insect Societies*. W.W. Norton & Company.

Huang, Y., Shen, X. J., Zou, Q., Wang, S. P., Tang, S. M. and Zhang, G. Z. (2011). Biological functions of microRNAs: A review. *J. Physiol. Biochem.* 67, 129–139.

Hughes, W. O. H., Oldroyd, B. P., Beekman, M. and Ratnieks, F. L. W. (2008). Ancestral monogamy shows kin selection is key to the evolution of eusociality. *Science*. 320, 1213-6.

Jones, P. A. and Takai, D. (2001). The role of DNA methylation in mammalian epigenetics. *Science* 293, 1068–1070.

Jones, B. M., Kingwell, C. J., Wcislo, W. T. and Robinson, G. E. (2017). Caste-biased gene expression in a facultatively eusocial bee suggests a role for genetic accommodation in the evolution of eusociality. *Proc. R. Soc. B Biol. Sci.* 284, 20162228.

Kapheim, K. M. and Johnson, M. M. (2017). Support for the reproductive ground plan hypothesis in a solitary bee: Links between sucrose response and reproductive status. *Proc. R. Soc. B Biol. Sci.* 284, 20162406.

Kapheim, K. M., Smith, A. R., Nonacs, P., Wcislo, W. T. and Wayne, R. K. (2013). Foundress polyphenism and the origins of eusociality in a facultatively eusocial sweat bee, *Megalopta genalis* (Halictidae). *Behav. Ecol. Sociobiol.* 67, 331–340.

Kapheim, K. M., Pan, H., Li, C., Salzberg, S. L., Puiu, D., Magoc, T., Robertson, H.M., Hudson, M.E., Venkat, A., Fischman, B.J. et al. (2015). Genomic signatures of evolutionary transitions from solitary to group living. *Science* 348, 1139–1143.

Kass, S. U., Pruss, D. and Wolffe, A. P. (1997). How does DNA methylation repress transcription? *Trends Genet.* 13, 444–449.

Khamis, A. M., Hamilton, A. R., Medvedeva, Y. A., Alam, T., Alam, I., Essack, M., Umylny, B., Jankovic, B. R., Naeger, N. L., Suzuki, M., et al. (2015). Insights into the transcriptional architecture of behavioral plasticity in the honey bee *Apis mellifera*. *Sci. Rep.* 5, 11136.

Kocher, S. D. and Paxton, R. J. (2014). Comparative methods offer powerful insights into social evolution in bees. *Apidologie* 45, 289–305.

Kocher, S. D., Mallarino, R., Rubin, B. E. R., Yu, D. W., Hoekstra, H. E. and Pierce, N. E. (2018). The genetic basis of a social polymorphism in halictid bees. *Nat. Commun.* 9, 4338.

Kornberg, R. D. and Lorch, Y. (1995). Interplay between chromatin structure and transcription. *Curr. Opin. Cell Biol.* 7, 371–375.

Kucharski, R., Maleszka, J., Foret, S. and Maleszka, R. (2008). Nutritional control of reproductive status in honeybees via DNA methylation. *Science* 319, 1827–1830.

Latchman, D. S. (1993). Transcription factors: an overview. *Int. J. Exp. Path* 29, 1305–1312.

Linksvayer, T. A. and Wade, M. J. (2005). The evolutionary origin and elaboration of sociality in the aculeate Hymenoptera: Maternal effects, sib-social effects, and heterochrony. *Q. Rev. Biol.* 80, 317–336.

Lu, F., Liu, Y., Inoue, A., Suzuki, T., Zhao, K. and Zhang, Y. (2016). Establishing chromatin regulatory landscape during mouse preimplantation development. *Cell* 165, 1375–1388.

Lyko, F., Foret, S., Kucharski, R., Wolf, S., Falckenhayn, C. and Maleszka, R. (2010). The honey bee epigenomes: Differential methylation of brain DNA in queens and workers. *PLoS Biol.* 8, e1000506.

Marshall, H., Lonsdale, Z. N. and Mallon, E. B. (2019). Methylation and gene expression differences between reproductive castes of bumblebee workers. *bioRxiv* 517698.

McNeill, M. S., Kapheim, K. M., Brockmann, A., McGill, T. A. W. and Robinson, G. E. (2016). Brain regions and molecular pathways responding to food reward type and value in honey bees. *Genes, Brain Behav.* 15, 305–317.

Michener, C. D. (1974). *The Social Behavior of the Bees: A Comparative Study*. Cambridge, MA: Harvard University Press.

Moczek, A. P., Sultan, S., Foster, S., Ledon-Rettig, C., Dworkin, I., Nijhout, H. F., Abouheif, E. and Pfennig, D. W. (2011). The role of developmental plasticity in evolutionary innovation. *Proc. R. Soc. B Biol. Sci.* 278, 2705-2713.

Moreau, C. S., Bell, C. D., Vila, R., Archibald, S. B. and Pierce, N. E. (2006). Phylogeny of the ants: Diversification in the age of angiosperms. *Science* 312, 101–104.

Naeger, N. L., Peso, M., Even, N., Barron, A. B. and Robinson, G. E. (2013). Altruistic behavior by egg-laying worker honeybees. *Curr. Biol.* 23, 1574–1578.

Narlikar, G. J., Fan, H. Y. and Kingston, R. E. (2002). Cooperation between complexes that regulate chromatin structure and transcription. *Cell* 108, 475–487.

Nelson, C. M., Ihle, K. E., Fondrk, M. K., Page, R. E. and Amdam, G. V. (2007). The gene vitellogenin has multiple coordinating effects on social organization. *PLoS Biol.* 5, e62.

O'Connor, V. and Lee, A. G. (2002). Synaptic vesicle fusion and synaptotagmin: 2B or not 2B? *Nat. Neurosci.* 5, 823-824.

Orphanides, G. and Reinberg, D. (2002). A unified theory of gene expression. *Cell* 108, 439–451.

Page Jr, R. E., Scheiner, R., Erber, J. and Amdam, G. V (2006). The development and evolution of division of labor and foraging specialization in a social insect (*Apis mellifera* L.). *Curr, Top. Dev. Biol.* 74, 253–286.

Page Jr, R. E., Rueppell, O. and Amdam, G. V (2012). Genetics of reproduction and regulation of honey bee (*Apis mellifera* L.) social behavior. *Annu. Rev. Genet.* 46, 19–97.

Pfennig, D. W., Wund, M. A., Snell-Rood, E. C., Cruickshank, T., Schlichting, C. D. and Moczek, A. P. (2010). Phenotypic plasticity's impacts on diversification and speciation. *Trends Ecol. Evol.* 25, 459–467.

Pigliucci, M. (2006). Phenotypic plasticity and evolution by genetic assimilation. *J. Exp. Biol.* 209, 2362–2367.

Raff, R. A. (2000). Evo-devo: The evolution of a new discipline. *Nat. Rev. Genet.* 1, 74-79.

Robinson, G. E. (1985). Effects of a juvenile hormone analogue on honey bee foraging behaviour and alarm pheromone production. *J. Insect Physiol.* 31, 277–282.

Romiguier, J., Cameron, S. A., Woodard, S. H., Fischman, B. J., Keller, L. and Praz, C. J. (2016). Phylogenomics controlling for base compositional bias reveals a single origin of eusociality in corbiculate bees. *Mol. Biol. Evol.* 33, 670–678.

Saul, M. C., Seward, C. H., Troy, J. M., Zhang, H., Sloofman, L. G., Lu, X., Weisner, P. A., Caetano-Anolles, D., Sun, H., Zhao, S. D., et al. (2017). Transcriptional regulatory dynamics drive coordinated metabolic and neural response to social challenge in mice. *Genome Res.* 27, 959-972.

Schulz, D. J., Sullivan, J. P. and Robinson, G. E. (2002). Juvenile hormone and octopamine in the regulation of division of labor in honey bee colonies. *Horm. Behav.* 42, 222–231.

Schwarz, M. P., Richards, M. H. and Danforth, B. N. (2007). Changing paradigms in insect social evolution: insights from halictine and allodapine bees. *Annu. Rev. Entomol.* 52, 127–150.

Shi, Y. Y., Zheng, H. J., Pan, Q. Z., Wang, Z. L. and Zeng, Z. J. (2015). Differentially expressed microRNAs between queen and worker larvae of the honey bee (*Apis mellifera*). *Apidologie* 46, 35–45.

Shi, Y. Y., Liu, H., Qiu, Y. F., Ma, Z. Y. and Zeng, Z. J. (2017). DNA methylation comparison between 4-day-old queen and worker larvae of honey bee. *J. Asia. Pac. Entomol.* 20, 299–303.

Shpigler, H. Y., Saul, M. C., Murdoch, E. E., Cash-Ahmed, A. C., Seward, C. H., Sloofman, L., Chandrasekaran, S., Sinha, S., Stubbs, L. J. and Robinson, G. E. (2017a). Behavioral, transcriptomic and epigenetic responses to social challenge in honey bees. *Genes, Brain Behav.* 16, 579–591.

Shpigler, H. Y., Saul, M. C., Corona, F., Block, L., Ahmed, A. C., Zhao, S. D. and Robinson, G. E. (2017b). Deep evolutionary conservation of autism-related genes. *Proc. Natl. Acad. Sci.* 114, 9653-9658.

Shpigler, H. Y., Saul, M. C., Murdoch, E. E., Corona, F., Cash-Ahmed, A. C., Seward, C. H., Chandrasekaran, S., Stubbs, L. J. and Robinson, G. E. (2018). Honey bee neurogenomic responses to affiliative and agonistic social interactions. *Genes, Brain Behav.* e12509.

Simola, D. F., Graham, R. J., Brady, C. M., Enzmann, B. L., Desplan, C., Ray, A., Zwiebel, L. J., Bonasio, R., Reinberg, D., Liebig, J., et al. (2015). Epigenetic (re)programming of caste-specific behavior in the ant *Camponotus floridanus*. *Science* 351, aac6633.

Smith, J. M. and Szathmaáry, E. (1995). *The Major Transitions in Evolution*. Oxford: Oxford University Press.

Snyder-Mackler, N., Sanz, J., Kohn, J. N., Voyles, T. N., Pique-Regi, R., Wilson, M. E., Barreiro, L. B. and Tung, J. (2018). Social status alters chromatin accessibility and the gene regulatory response to glucocorticoid stimulation in rhesus macaques. *PNAS* 116, 1219–1228.

Spitz, F. and Furlong, E. E. M. (2012). Transcription factors: From enhancer binding to developmental control. *Nat. Rev. Genet.* 13, 613-626.

Toth, A. L., Varala, K., Newman, T. C., Miguez, F. E., Hutchison, S. K., Willoughby, D. A., Simons, J. F., Egholm, M., Hunt, J. H., Hudson, M. E., et al. (2007). Wasp gene expression supports an evolutionary link between maternal behavior and eusociality. *Science* 318, 441–444.

Toth, A. L., Várala, K., Henshaw, M. T., Rodriguez-Zas, S. L., Hudson, M. E. and Robinson, G. E. (2010). Brain transcriptomic analysis in paper wasps identifies genes associated with behaviour across social insect lineages. *Proc. R. Soc. B Biol. Sci.* 277, 2139–2148.

Venkatesh, S. and Workman, J. L. (2015). Histone exchange, chromatin structure and the regulation of transcription. *Nat. Rev. Mol. cell Biol. Biol.* 16, 178-189.

Wang, Y., Kocher, S. D., Linksvayer, T. A., Grozinger, C. M., Page Jr, R. E. and Amdam, G. V (2011). Regulation of behaviorally associated gene networks in worker honey bee ovaries. *J. Exp. Biol.* 215, 124–134.

West-Eberhard, M. J. (1987). *Flexible strategy and social evolution*. (eds. Ito, Y., Brown, J. L., and Kikkawa, J.). Tokyo: Japan Scientific Societies Press.

West-Eberhard, M. J. (1996). Wasp societies as microcosms for the study of development and evolution. In *Natural history and evolution of paper wasps (eds Turillazzi S, West-Eberhard MJ)*, pp. 290–317.

West-Eberhard, M. J. (2003). *Developmental plasticity and evolution*. Oxford: Oxford University Press.

Wheeler, M. M., Ament, S. A., Rodriguez-Zas, S. L. and Robinson, G. E. (2013). Brain gene expression changes elicited by peripheral vitellogenin knockdown in the honey bee. *Insect Mol. Biol.* 22, 562–573.

Whitfield, C. W., Cziko, A. M. and Robinson, G. E. (2003). Gene expression profiles in the brain predict behavior in individual honey bees. *Science* 302, 296–299.

Whitfield, C. W., Ben-Shahar, Y., Brillet, C., Leoncini, I., Crauser, D., LeConte, Y., Rodriguez-Zas, S. and Robinson, G. E. (2006). Genomic dissection of behavioral maturation in the honey bee. *Proc. Natl. Acad. Sci.* 103, 16068–16075.

Wilson, E. O. (1971). *The insect societies.* Cambridge: Harvard University Press.

Wilson, E. O. (1980). Caste and division of labor in leaf-cutter ants (Hymenoptera: Formicidae: *Atta*). *Behav. Ecol. Sociobiol.* 7, 157–165.

Wojciechowski, M., Lowe, R., Maleszka, J., Conn, D., Maleszka, R. and Hurd, P. J. (2018). Phenotypically distinct female castes in honey bees are defined by alternative chromatin states during larval development. *Genome Res.* 28, 1532–1542.

Woodard, S. H., Fischman, B. J., Venkat, A., Hudson, M. E., Varala, K., Cameron, S. A., Clark, A. G. and Robinson, G. E. (2011). Genes involved in convergent evolution of eusociality in bees. *Proc. Natl. Acad. Sci.* 108, 7472–7477.

Xi, H., Shulha, H. P., Lin, J. M., Vales, T. R., Fu, Y., Bodine, D. M., McKay, R. D. G., Chenoweth, J. G., Tesar, P. J., Furey, T. S., et al. (2007). Identification and characterization of cell type-specific and ubiquitous chromatin regulatory structures in the human genome. *PLoS Genet.* 3, e136.

Zayed, A. and Robinson, G. E. (2012). Understanding the relationship between brain gene expression and social behavior: Lessons from the honey bee. *Annu. Rev. Genet.* 46, 591–615.

APPENDIX A: SUPPLEMENTARY FIGURES

Figure A.1

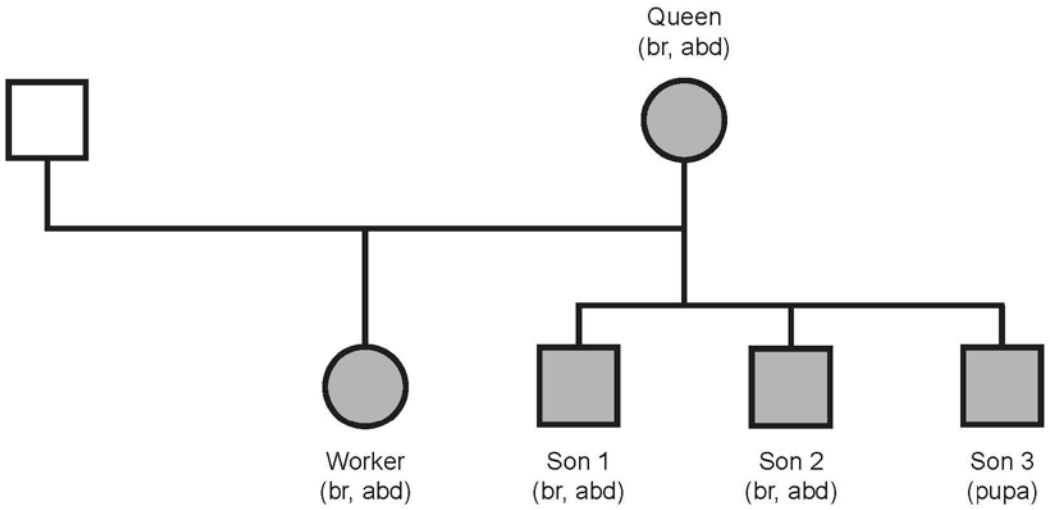


Figure A.1. Individuals used for transcriptome assembly.
Pedigree with the relationship of five individuals used in the transcriptome reference assembly (grayed shapes indicate individuals sequenced). Since four of the five individuals were adults (abdominal and brain tissues sequenced in separate libraries), a total of nine libraries were included in the assembly. br: brain, abd: abdomen

Figure A.2

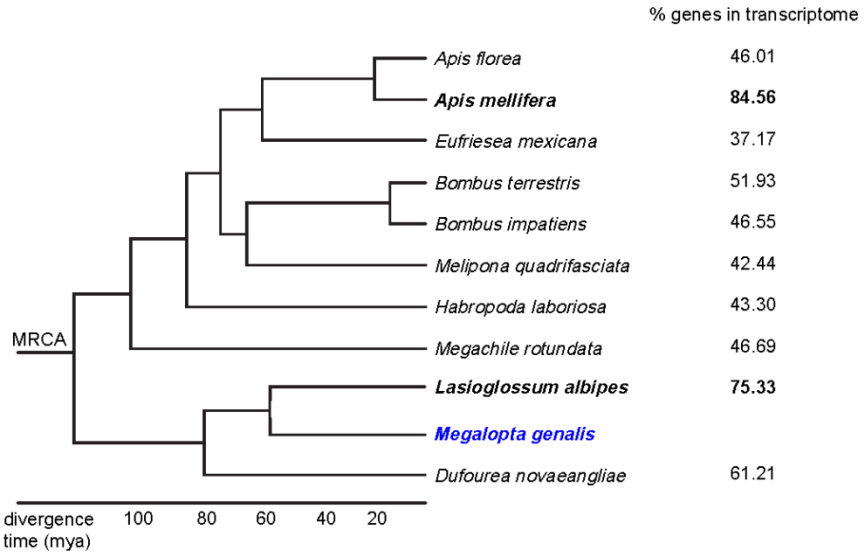


Figure A.2. Homology of contigs with 10 bee species.

Phylogeny of the 10 bees used for assessment of completeness of assembled contigs (including phylogenetic placement of *M. genalis*; tree and divergence times from Cardinal and Danforth (2013). The percentage of total genes within each species' genome with a BLASTN hit (e-value < 10-3) from at least one *M. genalis* contig is given to the right of the species name. The highest two percentages of gene lists with homologous *M. genalis* contigs are in bold.

Figure A.3

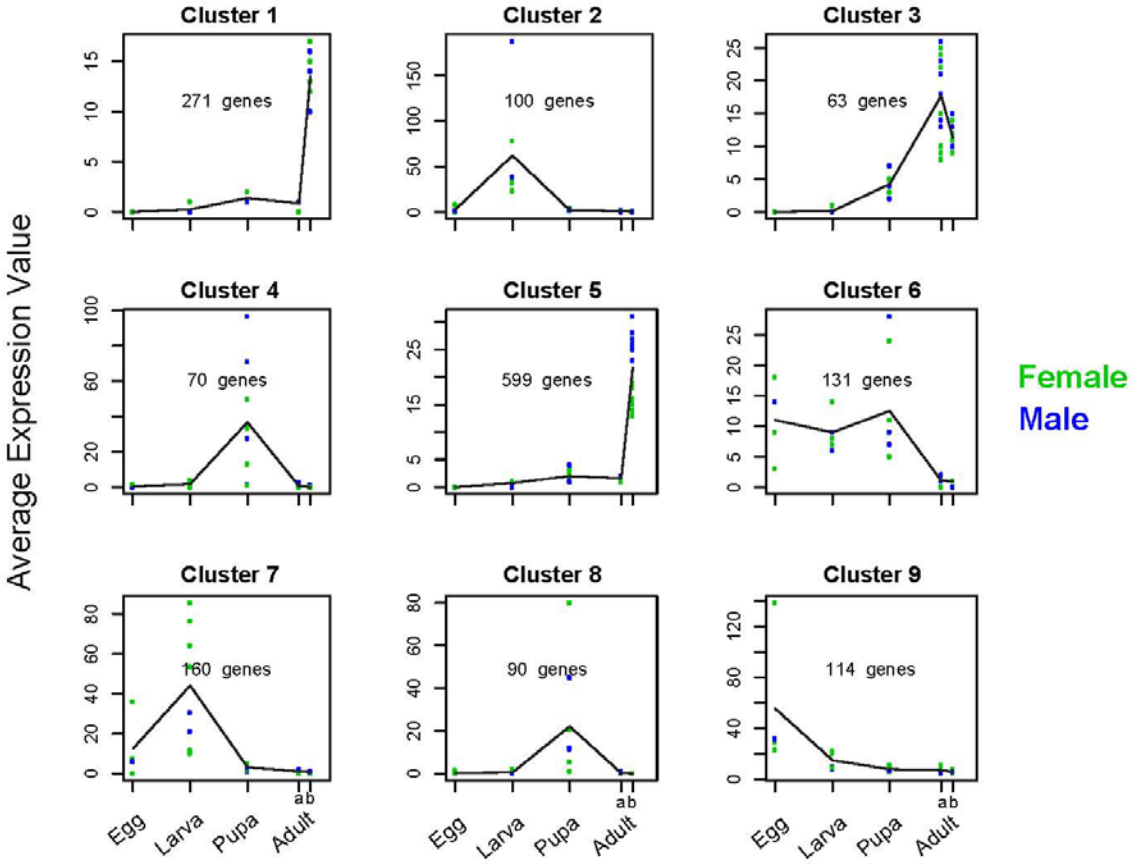


Figure A.3. Clustering analysis with 9 clusters.

Clusters of genes showing differences in expression associated with life stage, using the maximum allowable clusters (n=9). Each dot represents the median expression for all genes within the cluster for one individual, and lines connect the average expression value across individuals for each life stage. Along the x-axis, "a" and "b" refer to adult abdominal and brain tissues, respectively.

Figure A.4

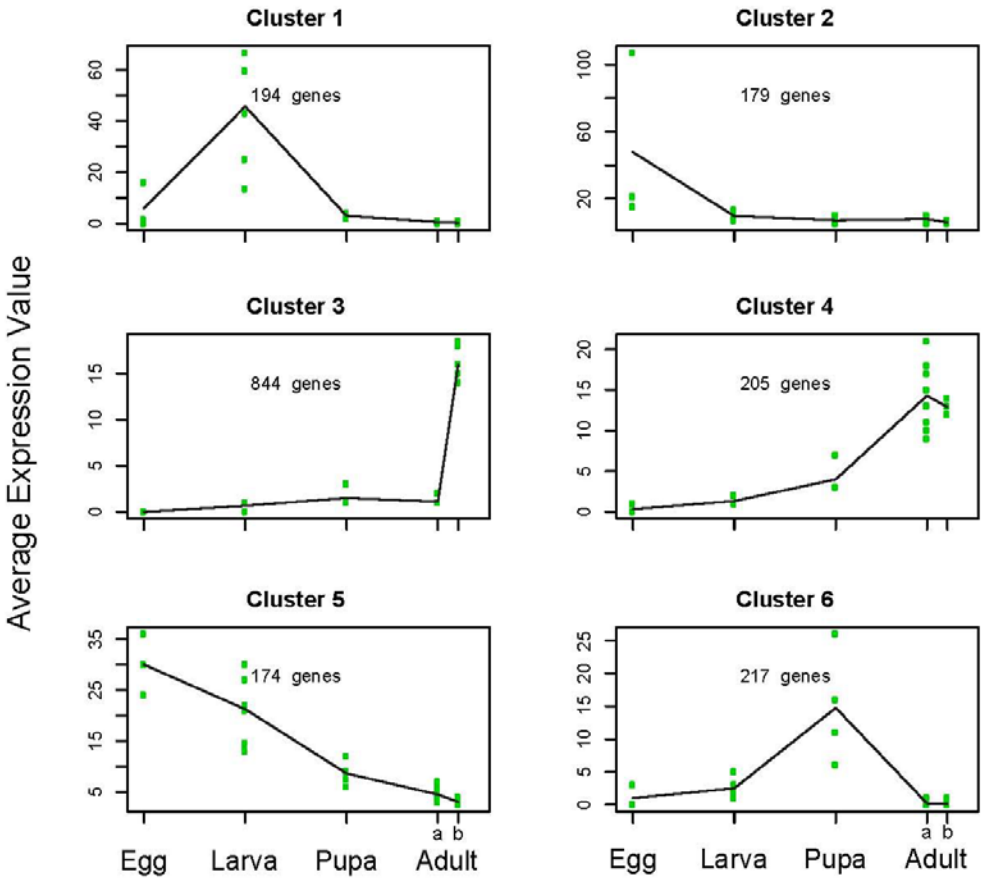


Figure A.4. Clustering analysis of females only.

Clusters of genes showing differences in expression associated with life stage excluding males. Each dot represents the median expression for all genes within the cluster for one individual female, and lines connect the average expression value across individuals for each life stage. Along the x-axis, “a” and “b” refer to adult abdominal and brain tissues, respectively.

Figure A.5

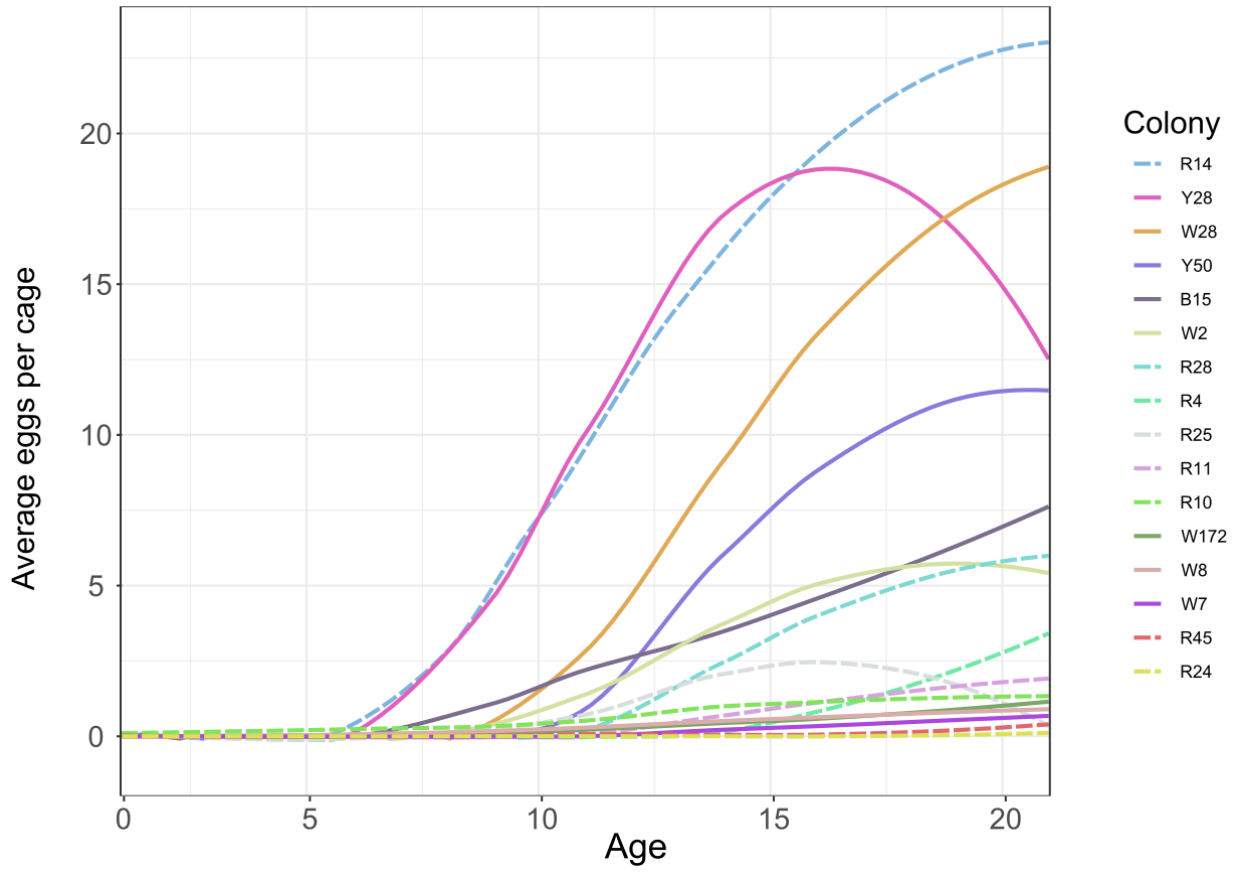


Figure A.5. Smoothed average egg counts for caged bees.

Solid lines indicate bees from source colonies headed by a naturally-mated queen, while dashed lines indicate bees from source colonies headed by a single-drone inseminated (SDI) queen.

Figure A.6

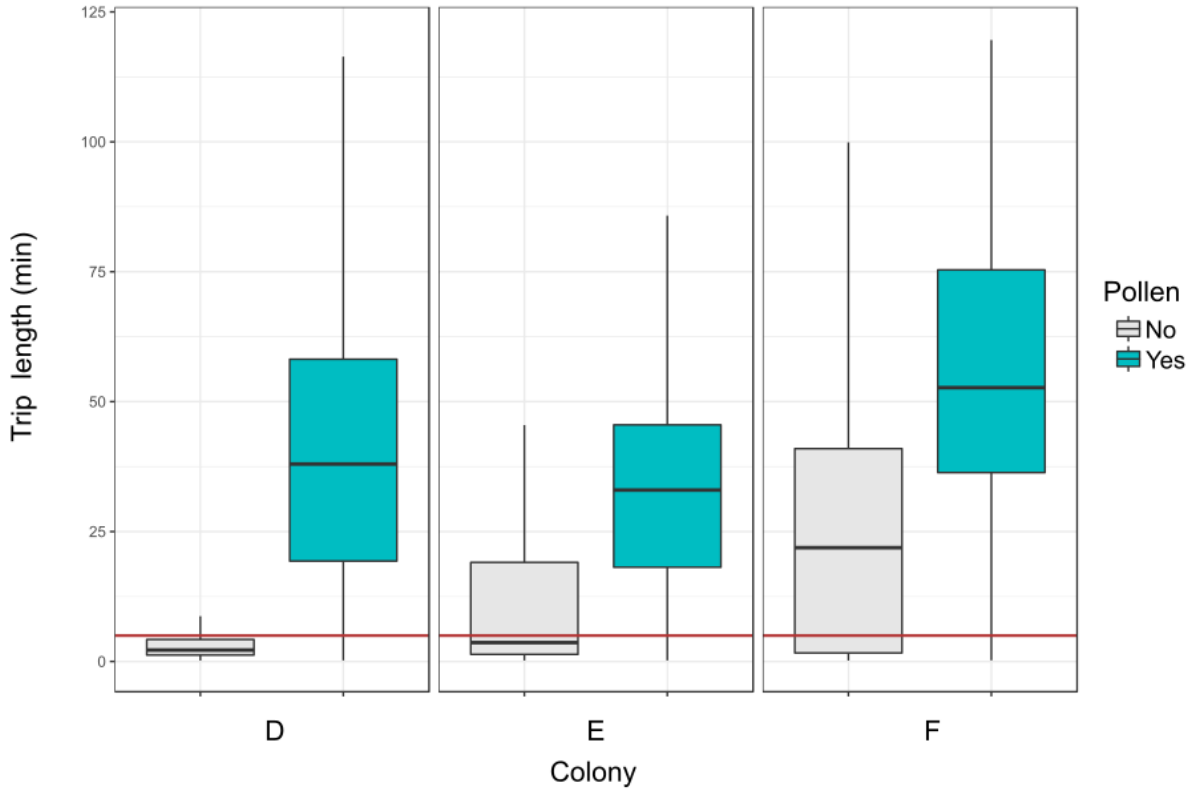
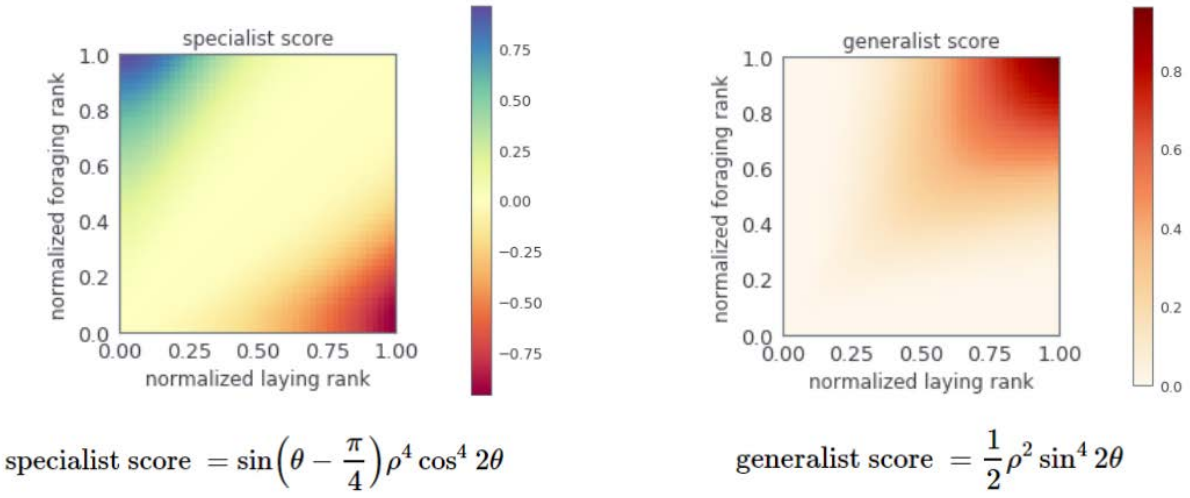


Figure A.6. Trip length separated by presence or absence of pollen on the hind legs of returning foragers based on manual annotation of incoming trips for colonies D-F. Horizontal red line indicates a trip length of 5 minutes. 7829 of 7925 (98.79%) of all trips annotated with pollen were longer than 5 minutes.

Figure A.7



$$\text{specialist score} = \sin\left(\theta - \frac{\pi}{4}\right)\rho^4 \cos^4 2\theta$$

$$\text{generalist score} = \frac{1}{2}\rho^2 \sin^4 2\theta$$

Figure A.7. Formulae and color-space mapping for specialist and generalist behavioral scores.

Figure A.8

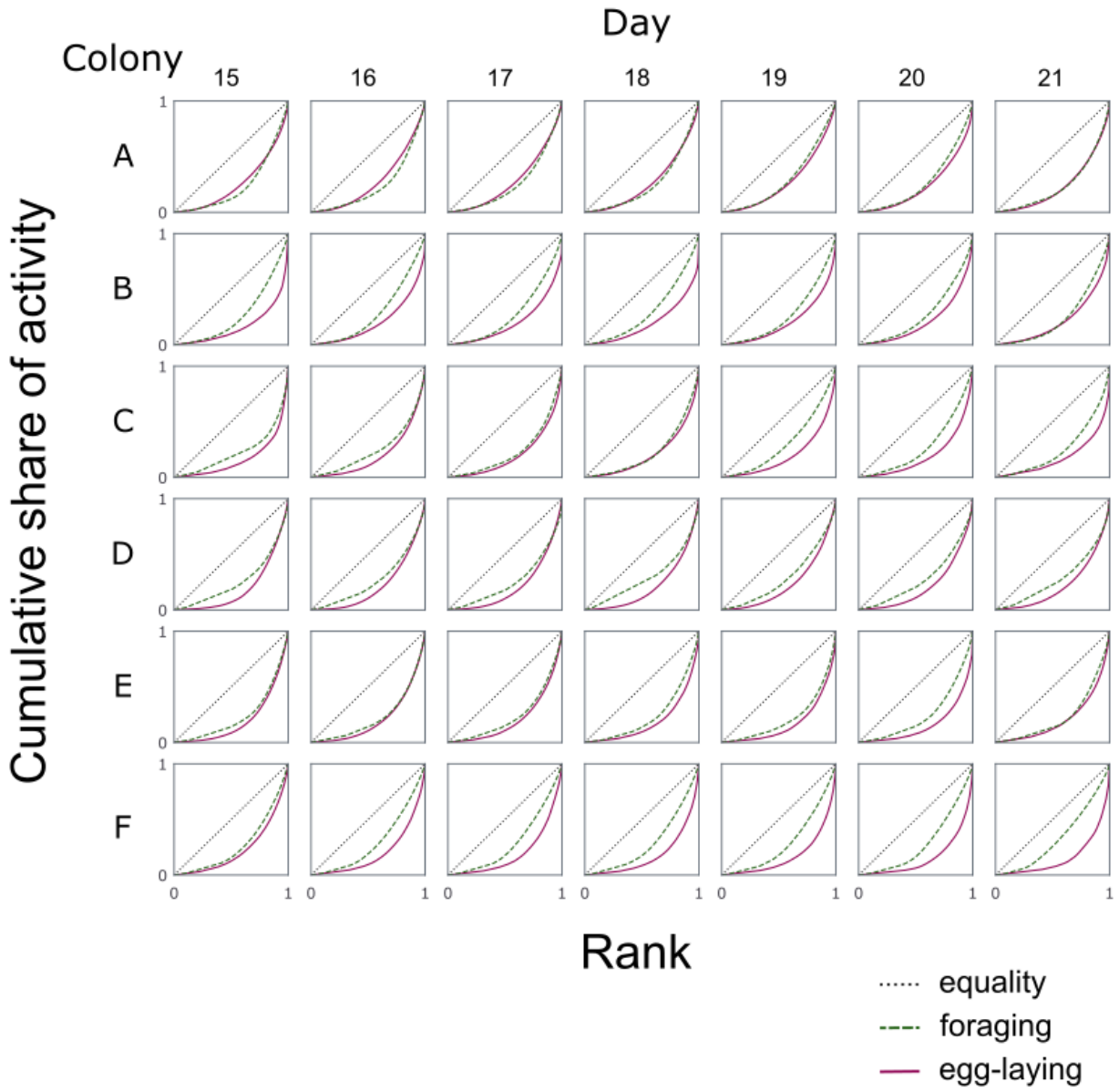


Figure A.8. Lorenz curve for all colonies and days. Individual bees were ranked by the number of foraging or egg-laying counts per day, and the fraction of each bee's contribution to the total activity of the colony was cumulatively plotted. If bees contributed equally to a given behavior, the curve would fall along the dotted line (labeled "equality"). Both foraging (green, dashed line) and egg-laying (purple, solid line) activities were unequally distributed among bees, similar to what was previously reported for foraging activity in queenright colonies (Tenczar et al., 2014).

Figure A.9

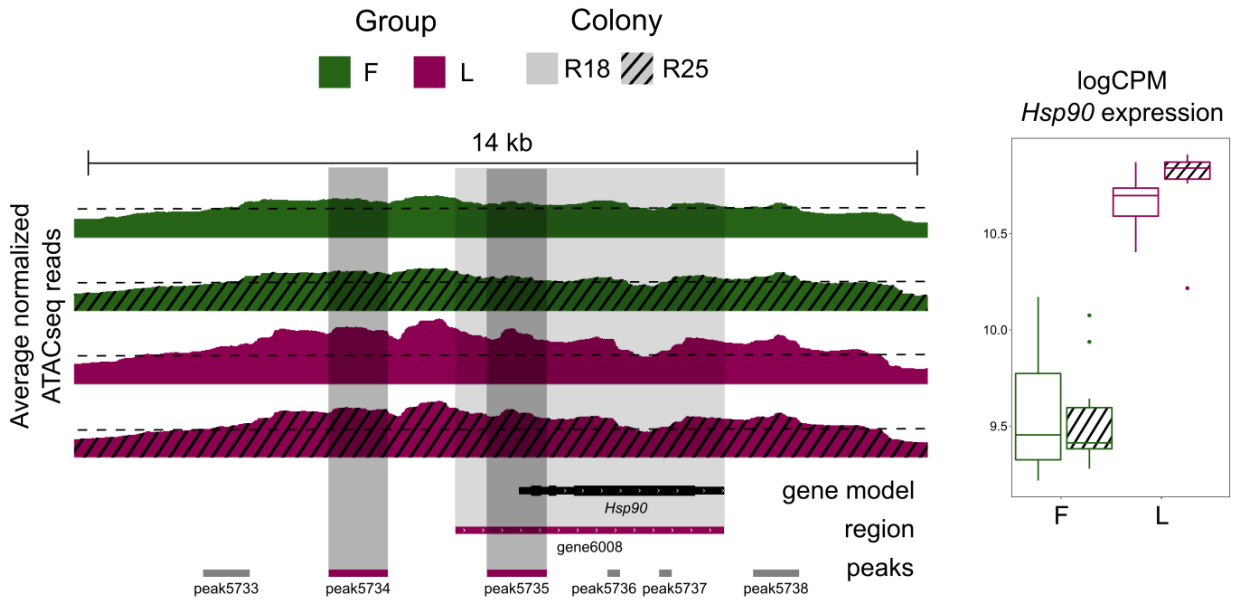


Figure A.9. Accessibility near *Hsp90*, a gene with differential accessibility (DAR: FDR-corrected p=0.000235, DAPs: FDR-corrected p=0.001854 for peak5734, 0.000465 for peak5735) and differential expression (DEG: FDR-corrected p=5.68e-12) between specialized foragers (green) and layers (purple). Color of the gene region and peaks indicates direction of differential accessibility; grey peaks are not differentially accessible. Shaded rectangles highlight differentially accessible peaks or regions. N=37 (6 R18 F, 15 R25 F, 8 R18 L, 8 R25 L).

Figure A.10

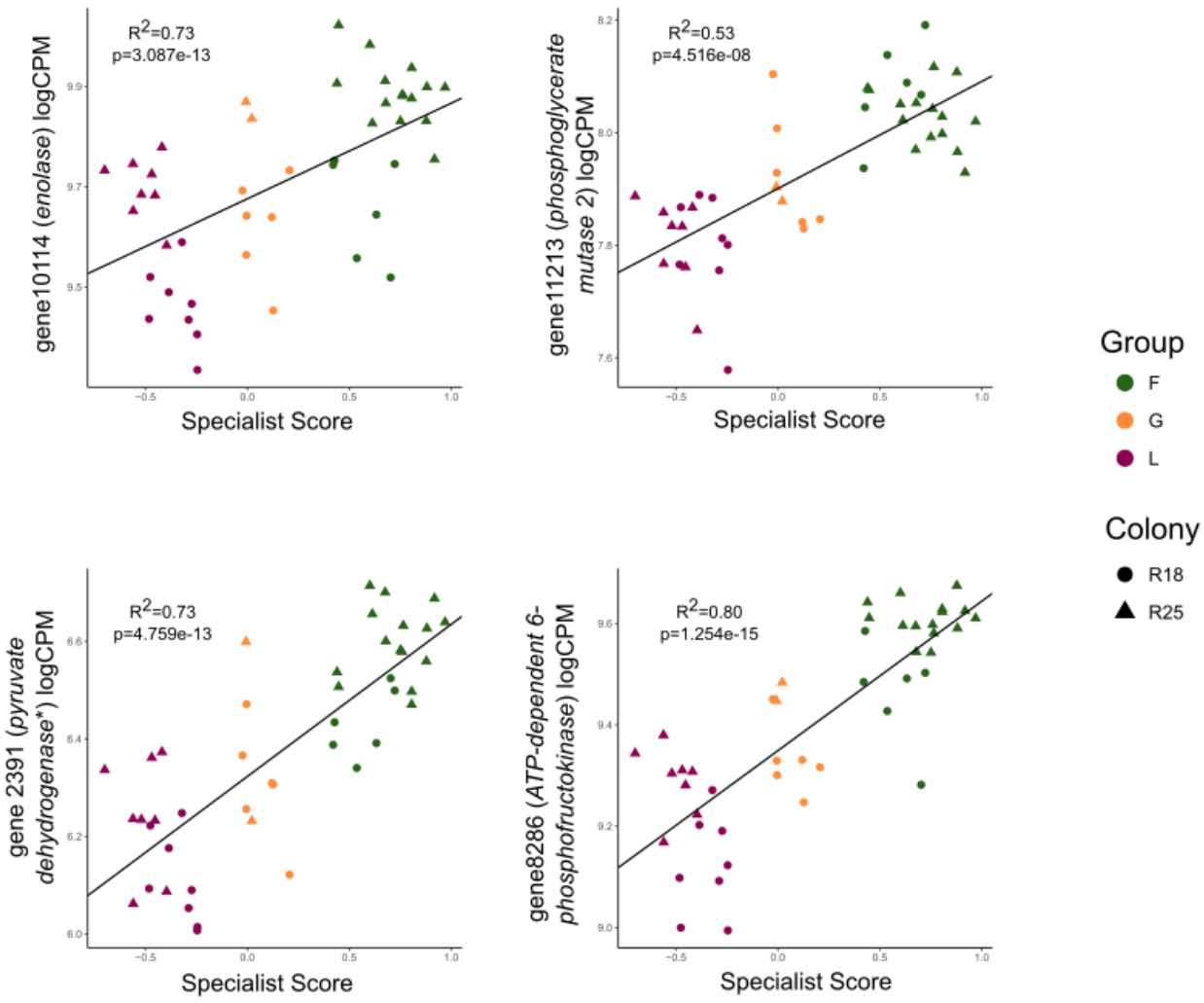


Figure A.10. Correlations between logCPM expression of four genes in the glycolysis/gluconeogenesis KEGG pathway and behavioral specialist score. These four genes were differentially expressed (FDR<0.05) between foragers (F) and layers (L) in the current study, as well as differentially expressed in at least three additional studies of caste-related behavioral phenotypes previously reported. Correlation coefficients and p-values are given for model which includes specialization score and colony. G: generalist. *full gene name of gene2391 is *pyruvate dehydrogenase E1 component subunit beta, mitochondrial.*

References

Cardinal, S., and Danforth, B.N. (2013). Bees diversified in the age of eudicots. *Proc. R. Soc. B Biol. Sci.* 280, 20122686.

Tenczar, P., Lutz, C. C., Rao, V. D., Goldenfeld, N. and Robinson, G. E. (2014). Automated monitoring reveals extreme interindividual variation and plasticity in honeybee foraging activity levels. *Anim. Behav.* 95, 41–48.

APPENDIX B: SUPPLEMENTARY TABLES

Table B.1. Source colony and dates for each experimental colony. A total of 800 bees were barcoded for each colony and tracked for 21 days. Colonies D-F included bees from two source colonies as indicated, with equal numbers of bees (400) from each.

Colony	Experiment ID	Source Colony (Colonies)	Dates
A	2016-03	W2 (naturally mated)	06/26/16-07/17/16
B	2016-02	W2 (naturally mated)	06/03/16-06/23/16
C	2016-08	W28 (naturally mated)	08/21/16-09/10/16
D	2016-06	R18 (single drone inseminated, SDI) and R45 (SDI)	08/02/16-08/22/16
E	2016-04	R18 (SDI) and R45 (SDI)	07/07/16-07/27/16
F	2016-05	R12 (SDI) and R25 (SDI)	07/14/16-08/04/16

Table B.2. Ovary and behavioral information for select bees dissected from colonies E and F. Specialist Score and Generalist Score are median scores over the final 7 days of tracking (days 14-21). Ovary Score is on a scale of 1-5, where 1-2 are undeveloped and 3-5 are developed to different degrees (i.e., 1 least developed, 5 most developed). Numbers of egg-laying (L), foraging (F), and trophallaxis (T) events are total counts for each bee for the final 7 days of tracking.

Bee	Colony	Specialist Score	Generalist Score	Ovary Score (1-2: undeveloped; 3-5: developed)	# Ovar-ioles	# L	# F	# T
E_1793	R45	1.000	0.000	2	8	0	333	568
F_247	R25	0.990	0.000	1	-	0	139	428
E_1634	R45	0.984	0.000	2	8	0	175	360
F_255	R25	0.974	0.000	2	6	0	127	375
E_1747	R45	0.973	0.000	1	9	0	197	434
F_1096	R25	0.953	0.000	3	25	0	130	302
E_1756	R45	0.952	0.000	2	8	0	196	637
E_1752	R45	0.937	0.000	1	8	0	129	370
E_1646	R45	0.912	0.000	1	5	1	118	315
F_1077	R25	0.903	0.000	4	11	0	110	313
F_907	R25	0.884	0.000	4	9	1	106	417
E_1687	R45	0.880	0.000	1	6	0	144	617
F_277	R25	0.868	0.000	4	11	0	117	332
F_202	R25	0.844	0.000	5	67	0	100	320
F_857	R25	0.825	0.000	4	50	0	110	341
F_54	R25	0.821	0.000	4	13	1	83	452
E_110	R18	0.793	0.000	3	9	1	101	264
E_1968	R45	0.757	0.060	2	6	2	221	490
E_1355	R45	0.746	0.000	3	9	1	116	626
E_77	R18	0.733	0.060	3	9	2	149	327
E_1984	R45	0.729	0.061	1	6	2	133	398
F_245	R25	0.719	0.000	5	48	0	91	330
F_875	R25	0.686	0.069	4	23	2	88	258
F_157	R25	0.664	0.070	4	24	2	100	350
E_331	R18	0.606	0.065	2	6	2	95	369
F_348	R25	0.591	0.000	1	5	0	79	289
F_354	R25	0.591	0.000	2	5	0	82	339
E_205	R18	0.579	0.132	3	5	3	163	440
E_1413	R45	0.575	0.132	3	6	3	191	508
E_52	R18	0.558	0.133	2	4	3	124	322
F_1048	R25	0.520	0.191	4	45	3	158	436

(Table B.2 continued)

E_173	R18	0.395	0.278	3	7	5	150	261
E_1181	R18	0.383	0.294	2	4	7	172	187
E_144	R18	0.360	0.240	3	5	5	99	376
F_1015	R25	0.210	0.377	3	12	5	97	426
E_725	R18	0.208	0.374	2	4	7	111	472
E_107	R18	0.207	0.430	2	9	9	108	443
E_185	R18	0.207	0.325	1	7	7	77	387
F_1098	R25	0.161	0.388	3	3	4	99	306
E_1892	R45	0.093	0.584	3	7	19	137	521
E_189	R18	0.088	0.598	3	5	20	180	474
F_316	R25	0.000	0.881	5	9	58	86	609
F_193	R25	0.000	0.981	4	5	321	94	796
E_138	R18	-0.001	0.898	3	6	151	85	421
E_208	R18	-0.001	0.835	4	8	87	65	565
E_921	R18	-0.002	0.904	4	7	209	73	713
E_14	R18	-0.004	0.778	4	7	90	42	686
E_305	R18	-0.010	0.839	4	7	198	59	863
E_85	R18	-0.011	0.829	4	7	192	51	1096
E_55	R18	-0.066	0.671	5	12	215	27	667
E_529	R18	-0.099	0.625	5	8	287	19	915
E_862	R18	-0.162	0.429	5	8	102	9	457
E_330	R18	-0.318	0.290	5	7	137	5	712
E_66	R18	-0.336	0.287	4	7	160	6	559
F_1137	R25	-0.425	0.214	5	17	112	20	617
F_398	R25	-0.472	0.209	5	13	189	17	652
F_968	R25	-0.494	0.207	4	36	360	10	467
F_904	R25	-0.624	0.083	5	19	107	6	516
E_182	R18	-0.714	0.000	5	10	84	8	523
F_246	R25	-0.719	0.079	5	95	763	10	774
E_129	R18	-0.849	0.000	5	14	134	11	806
F_1053	R25	-0.866	0.000	4	20	96	10	334
E_285	R18	-0.916	0.000	4	9	190	13	645
F_212	R25	-0.971	0.000	5	22	204	12	664
F_206	R25	-0.979	0.000	5	37	316	8	585

Table B.3. Ovary and behavioral information for sequenced bees. Group indicates whether bee was categorized as a forager (F), generalist (G), or layer (L) based on overall behavior. Specialist Score and Generalist Score are median scores over the final 7 days of tracking (days 14-21). Ovary Score is on a scale of 1-5, where 1-2 are undeveloped and 3-5 are developed to different degrees (i.e., 1 least developed, 5 most developed). Numbers of egg-laying (L) and foraging (F) events are total counts for each bee for the final 7 days of tracking.

Bee	Group	Colony	Specialist Score	Generalist Score	Ovary Score (1-2: undeveloped; 3-5: developed)	# Ovarioles	# L	# F
F_247	F	R25	0.990	0.000	1	-	0	139
F_255	F	R25	0.974	0.000	2	6	0	127
F_1096	F	R25	0.953	0.000	3	25	0	130
F_1077	F	R25	0.903	0.000	4	11	0	110
F_907	F	R25	0.884	0.000	4	9	1	106
F_277	F	R25	0.868	0.000	4	11	0	117
F_202	F	R25	0.844	0.000	5	67	0	100
F_857	F	R25	0.825	0.000	4	50	0	110
F_54	F	R25	0.821	0.000	4	13	1	83
F_245	F	R25	0.719	0.000	5	48	0	91
F_157	F	R25	0.664	0.070	4	24	2	100
F_354	F	R25	0.591	0.000	2	5	0	82
E_205	F	R18	0.579	0.132	3	5	3	163
F_1048	F	R25	0.520	0.191	4	45	3	158
E_1181	F	R18	0.383	0.294	2	4	7	172
F_1015	F	R25	0.210	0.377	3	12	5	97
E_725	F	R18	0.208	0.374	2	4	7	111
E_107	F	R18	0.207	0.430	2	9	9	108
E_185	F	R18	0.207	0.325	1	7	7	77
F_1098	F	R25	0.161	0.388	3	3	4	99
E_189	F	R18	0.088	0.598	3	5	20	180
F_316	G	R25	0.000	0.881	5	9	58	86
F_193	G	R25	0.000	0.981	4	5	321	94
E_138	G	R18	-0.001	0.898	3	6	151	85
E_208	G	R18	-0.001	0.835	4	8	87	65
E_921	G	R18	-0.002	0.904	4	7	209	73
E_14	G	R18	-0.004	0.778	4	7	90	42
E_305	G	R18	-0.010	0.839	4	7	198	59
E_85	G	R18	-0.011	0.829	4	7	192	51

(Table B.3 continued)

E_55	L	R18	-0.066	0.671	5	12	215	27
E_529	L	R18	-0.099	0.625	5	8	287	19
E_862	L	R18	-0.162	0.429	5	8	102	9
E_330	L	R18	-0.318	0.290	5	7	137	5
E_66	L	R18	-0.336	0.287	4	7	160	6
F_1137	L	R25	-0.425	0.214	5	17	112	20
F_398	L	R25	-0.472	0.209	5	13	189	17
F_968	L	R25	-0.494	0.207	4	36	360	10
F_904	L	R25	-0.624	0.083	5	19	107	6
E_182	L	R18	-0.714	0.000	5	10	84	8
F_246	L	R25	-0.719	0.079	5	95	763	10
E_129	L	R18	-0.849	0.000	5	14	134	11
F_1053	L	R25	-0.866	0.000	4	20	96	10
E_285	L	R18	-0.916	0.000	4	9	190	13
F_212	L	R25	-0.971	0.000	5	22	204	12
F_206	L	R25	-0.979	0.000	5	37	316	8

Table B.4. Thirty-six genes which were differentially expressed between foragers and layers that were also implicated in at least 4 other studies of either caste (Alaux et al., 2009; Grozinger et al., 2007; Jones et al., 2017; Marshall et al., 2019; Toth et al., 2010; Wheeler et al., 2013; Whitfield et al., 2003) or molecular evolution of eusocial lineages (Harpur et al., 2014; Kapheim et al., 2015; Woodard et al., 2011). *A. mellifera* Q vs. W: Grozinger et al. 2007; *A. mellifera* N vs. F (1): Whitfield et al. 2003; *A. mellifera* N vs. F (2): Alaux et al. 2009; *A. mellifera Vg*RNAi: Wheeler et al. 2013; *B. terrestris* RW vs. W: Marshall et al. 2019; *M. genalis* Q vs. W: Jones et al. 2017; *P. metricus* Q vs. W: Toth et al. 2010; *A. mellifera* selection: Harpur et al. 2014; Selection in eusocial bees (1): Woodard et al. 2011; Selection in eusocial bees (2): Kapheim et al. 2015. Q: queen, W: worker, N: nurse, F: forager, *Vg*RNAi: *Vitellogenin* RNAi, RW: reproductive worker.

(Table B.4 continued)

	A. mellifera F vs. L (this study)	*A. mellifera* Q vs. W	*A. mellifera* N vs. F (1)	*A. mellifera* N vs. F (2)	*A. mellifera Vg* RNAi	*B. terrestris* RW vs. W	*M. genalis* Q vs. W	*P. metricus* Q vs. W	*A. mellifera* selection	Selection in eusocial bees (1)	Selection in eusocial bees (2)	Gene Description	KEGG Pathway(s)	Brite KEGG orthology (top hierchical term)	Brite KEGG orthology (second hierchical term)
gene4879	X	X	X	X	X				X			alpha,alpha-trehalose-phosphate synthase [UDP-forming]	starch and sucrose metabolism; metabolic pathways	metabolism	carbohydrate metabolism
gene426	X	X	X	X	X					X		ankyrin repeat domain-containing protein 12		genetic information processing	chromosome and associated proteins
gene10114	X	X	X		X					X	X	enolase	glycolysis/gluconeogenesis; Metabolic pathways; Carbon metabolism; Biosynthesis of amino acids; RNA degradation	metabolism	carbohydrate metabolism
gene11213	X	X	X		X					X	X	phosphoglycerate mutase 2	glycolysis/gluconeogenesis; Glycine, serine and threonine metabolism; Metabolic pathways; Carbon metabolism; Biosynthesis of amino acids	metabolism	carbohydrate metabolism
gene10032	X	X	X	X	X							transcription factor cwo		genetic information processing	transcription factors
gene10184	X	X	X	X	X							endoplasmic reticulum resident protein 44		genetic information processing	membrane trafficking
gene10225	X	X	X	X	X							protein N-terminal glutamine amidohydrolase		unclassified: metabolism	enzymes with EC numbers
gene1379	X	X	X	X	X							methionine sulphoxide reductase A		unclassified: metabolism	enzymes with EC numbers

(Table B.4 continued)

gene2132	X	X	X	X	X							uncharacterized protein LOC552101			
gene2387	X	X	X	X	X							inositol oxygenase	Ascorbate and aldarate metabolism; inositol phosphate metabolism	metabolism	carbohydrate metabolism
gene2466	X	X	X	X	X							putative inorganic phosphate cotransporter		signaling and cellular processes	transporters
gene5343	X	X	X	X	X							nucleosome assembly protein 1-like 1		genetic information processing	chromosome and associated proteins
gene5716	X	X	X	X	X							acidic phospholipase A2 PA4			
gene6012	X	X	X	X	X							glucose transporter type 1		signaling and cellular processe	transporters
gene8095	X	X	X	X	X							purine nucleoside phosphorylase	purine metabolism; pyrimidine metabolism; nicotinate and nicotinamide metabolism; metabolic pathways	metabolism	nucleotide metabolism
gene7280	X	X	X	X		X						methylthioribose-1-phosphate isomerase	cysteine and methionine metabolism, metabolic pathways	metabolism	amino acid metabolism
gene9114	X	X	X	X				X				putative glutathione-specific gamma-glutamylcyclotransferase 2	glutathione metabolism; metabolic pathways	metabolism	metabolism of other amino acids
gene7116	X	X	X	X					X			uncharacterized protein LOC100577363			
gene9985	X	X	X		X	X						6-phosphogluconate dehydrogenase, decarboxylating	pentose phosphate pathway; glutathione metabolism; metabolic pathways; carbon metabolism	metabolism	carbohydrate metabolism
gene11952	X	X	X		X		X					cathepsin L1	autophagy - animal; lysosome; phagosome	cellular processes	transport and catabolism
gene8888	X	X	X		X		X					neural/ectodermal development factor IMP-L2			
gene7198	X	X	X		X			X				H(+)/Cl(-) exchange transporter 5		signaling and cellular processes	ion channels
gene10407	X	X	X		X				X			neprilysin-4		metabolism	peptidases

(Table B.4 continued)

gene7883	X	X	X		X				X			40S ribosomal protein S12, mitochondrial	ribosome	genetic Information Processing	translation
gene2391	X	X	X		X					X		pyruvate dehydrogenase E1 component subunit beta, mitochondrial	glycolysis/gluconeogenesis; citrate cycle (TCA cycle); pyruvate metabolism; metabolic pathways; carbon metabolism	metabolism	carbohydrate metabolism
gene2111	X	X		X	X	X						acetyl-CoA acetyltransferase, cytosolic	fatty acid degradation; synthesis and degradation of ketone bodies; valine, leucine and isoleucine degradation; lysine degradation; tryptophan metabolism; pyruvate metabolism; glyoxylate and dicarboxylate metabolism; propanoate metabolism; butanoate metabolism; terpenoid backbone biosynthesis; metabolic pathways; carbon metabolism; fatty acid metabolism	metabolism	carbohydrate metabolism
gene10435	X	X		X	X				X			peptidoglycan-recognition protein S2 precursor		unclassified: signaling and cellular processes	structural proteins
gene2192	X	X			X	X			X			arylsulfatase B			
gene7330	X	X			X	X			X			long-chain-fatty-acid--CoA ligase 4	fatty acid biosynthesis; fatty acid degradation; metabolic pathways; fatty acid metabolism; peroxisome	metabolism	lipid metabolism
gene8675	X	X			X	X			X			uncharacterized protein LOC724286			

(Table B.4 continued)

gene9987	X	X			X					X	X	probable citrate synthase 2, mitochondrial	citrate cycle (TCA cycle); glyoxylate and dicarboxylate metabolism; metabolic pathways; carbon metabolism; 2-oxocarboxylic acid metabolism; biosynthesis of amino acids	metabolism	carbohydrate metabolism
gene11270	X	X						X	X	X		arginine kinase isoform X1	arginine and proline metabolism; metabolic pathways	metabolism	amino acid metabolism
gene8359	X		X	X	X				X			dual 3',5'-cyclic-AMP and -GMP phosphodiesterase 11			
gene8286	X		X		X				X	X		ATP-dependent 6-phosphofructokinase	glycolysisg/gluconeogenesis; pentose phosphate pathway; fructose and mannose metabolism; galactose metabolism; metabolic pathways; carbon metabolism; biosynthesis of amino acids; RNA degradation	metabolism	carbohydrate metabolism
gene11183	X			X	X			X		X		hexokinase type 2	glycolysis/gluconeogenesis; fructose and mannose metabolism; galactose metabolism; starch and sucrose metabolism; amino sugar and nucleotide sugar metabolism; metabolic pathways; carbon metabolism	metabolism	carbohydrate metabolism
gene4466	X				X				X	X	X	putative 28S ribosomal protein S5, mitochondrial isoform X2	ribosome	genetic Information Processing	translation

Table B.5. Genes both differentially expressed and with differential chromatin accessibility (based on both region- and peak-based approaches) between foragers (F) and layers (L). F>L indicates higher expression or accessibility in foragers relative to layers, L>F indicates higher expression or accessibility in layers relative to foragers. DEG: differentially expressed gene direction, DAR: differentially accessible region direction, DAP: differentially accessible peak direction.

gene	DEG	DAR	DAP	gene description
gene1026	F>L	F>L	F>L	protein Skeletor, isoforms B/C
gene10889	F>L	F>L	F>L	uncharacterized LOC410506
gene10900	F>L	F>L	F>L	uncharacterized LOC410514
gene11497	F>L	F>L	F>L	fibrillin-2
gene11505	F>L	F>L	F>L	uncharacterized LOC406065 precursor
gene1156	F>L	F>L	F>L	homeobox protein aristaless
gene11575	F>L	F>L	F>L	odorant binding protein 9
gene11663	F>L	F>L	F>L	multidrug resistance-associated protein 4
gene1205	F>L	F>L	F>L	facilitated trehalose transporter Tret1
gene12234	F>L	F>L	F>L	scavenger receptor class B member 1
gene1684	F>L	F>L	F>L	uncharacterized LOC724275
gene2229	F>L	F>L	F>L	CD151 antigen
gene2377	F>L	F>L	F>L	uncharacterized LOC410793
gene2598	F>L	F>L	F>L	uncharacterized LOC408759
gene3091	F>L	F>L	F>L	wiskott-Aldrich syndrome protein family member 3
gene3092	F>L	F>L	F>L	uncharacterized LOC102654960
gene315	F>L	F>L	F>L	uncharacterized LOC107966071
gene347	F>L	F>L	F>L	matrix metalloproteinase-2
gene3888	F>L	F>L	F>L	E3 ubiquitin-protein ligase MYLIP
gene4170	F>L	F>L	F>L	monocarboxylate transporter 14
gene4542	F>L	F>L	F>L	uncharacterized LOC113218812
gene5405	F>L	F>L	F>L	A disintegrin and metalloproteinase with thrombospondin motifs 7
gene5440	F>L	F>L	F>L	COUP transcription factor 2
gene5621	F>L	F>L	F>L	glutamine synthetase
gene6717	F>L	F>L	F>L	dnaJ homolog subfamily C member 11
gene68	F>L	F>L	F>L	uncharacterized LOC100576537
gene7051	F>L	F>L	F>L	Lachesin
gene8410	F>L	F>L	F>L	uncharacterized LOC102655465
gene8412	F>L	F>L	F>L	uncharacterized LOC102655526
gene8413	F>L	F>L	F>L	tyrosine-protein kinase Dnt
gene8948	F>L	F>L	F>L	uncharacterized LOC100576913
gene1002	L>F	L>F	L>F	histone H3.3-like type 1
gene10194	L>F	L>F	L>F	partner of Y14 and mago
gene10204	L>F	L>F	L>F	polycomb group protein Psc

(Table B.5 continued)

gene10225	L>F	L>F	L>F	protein N-terminal glutamine amidohydrolase
gene10345	L>F	L>F	L>F	uncharacterized LOC102656433
gene10447	L>F	L>F	L>F	Trichohyalin
gene10497	L>F	L>F	L>F	PI-PLC X domain-containing protein 3
gene10543	L>F	L>F	L>F	testis-specific serine/threonine-protein kinase 3
gene1068	L>F	L>F	L>F	short coiled-coil protein B
gene1069	L>F	L>F	L>F	fibroblast growth factor 1
gene12066	L>F	L>F	L>F	tektin-4
gene12182	L>F	L>F	L>F	uncharacterized LOC551079
gene1359	L>F	L>F	L>F	NTF2-related export protein
gene1739	L>F	L>F	L>F	protein lethal(2)essential for life
gene1772	L>F	L>F	L>F	L-galactose dehydrogenase
gene1916	L>F	L>F	L>F	uncharacterized LOC107966102
gene2021	L>F	L>F	L>F	Golgi membrane protein 1
gene2101	L>F	L>F	L>F	uncharacterized LOC102653599
gene2605	L>F	L>F	L>F	estradiol 17-beta-dehydrogenase 8
gene2611	L>F	L>F	L>F	heat shock protein cognate 3 precursor
gene2803	L>F	L>F	L>F	uncharacterized protein LOC100577047
gene2979	L>F	L>F	L>F	nuclear protein 1
gene300	L>F	L>F	L>F	intraflagellar transport protein 52 homolog
gene328	L>F	L>F	L>F	heat shock protein cognate 4
gene3560	L>F	L>F	L>F	viral IAP-associated factor homolog
gene3681	L>F	L>F	L>F	dnaJ homolog subfamily A member 1
gene3769	L>F	L>F	L>F	protein 4.1 homolog
gene4082	L>F	L>F	L>F	cyclin-dependent kinase-like 4
gene4267	L>F	L>F	L>F	WD repeat-containing protein 92
gene4615	L>F	L>F	L>F	estrogen-related receptor
gene464	L>F	L>F	L>F	arrestin domain-containing protein 17
gene4965	L>F	L>F	L>F	uncharacterized protein YER152C
gene5142	L>F	L>F	L>F	neurotrimin
gene5145	L>F	L>F	L>F	3-ketodihydrosphingosine reductase
gene5150	L>F	L>F	L>F	methyl-CpG-binding domain protein 2
gene5486	L>F	L>F	L>F	coiled-coil domain-containing protein 134
gene5650	L>F	L>F	L>F	methyltransferase-like protein 5
gene5933	L>F	L>F	L>F	U1 small nuclear ribonucleoprotein C
gene5939	L>F	L>F	L>F	acidic leucine-rich nuclear phosphoprotein 32 family member A
gene5956	L>F	L>F	L>F	histone H2A.V
gene6008	L>F	L>F	L>F	heat shock protein 90
gene6086	L>F	L>F	L>F	peptidoglycan-recognition protein LC
gene6580	L>F	L>F	L>F	mitotic spindle assembly checkpoint protein MAD1
gene7174	L>F	L>F	L>F	RNA-binding protein spenito
gene7625	L>F	L>F	L>F	uncharacterized LOC724178
gene7869	L>F	L>F	L>F	protein LSM14 homolog A

(Table B.5 continued)

gene8066	L>F	L>F	L>F	serine/threonine-protein phosphatase 1 regulatory subunit 10
gene810	L>F	L>F	L>F	tetratricopeptide repeat protein 4
gene8184	L>F	L>F	L>F	fasciculation and elongation protein zeta-2
gene8273	L>F	L>F	L>F	uncharacterized LOC113219030
gene8287	L>F	L>F	L>F	14-3-3 protein zeta
gene8706	L>F	L>F	L>F	uncharacterized LOC409534
gene8711	L>F	L>F	L>F	uncharacterized LOC409323
gene9622	L>F	L>F	L>F	ras suppressor protein 1
gene10605	F>L	F>L	L>F	neuropeptide-like 1
gene5102	F>L	F>L	L>F	calcium-dependent protein kinase 4
gene10184	F>L	L>F	L>F	endoplasmic reticulum resident protein 44
gene1039	F>L	L>F	L>F	histone H2A
gene1041	F>L	L>F	L>F	histone H1
gene10633	F>L	L>F	L>F	RNA-binding protein 40
gene10956	F>L	L>F	L>F	gamma-aminobutyric acid receptor-associated protein
gene1408	F>L	L>F	L>F	growth hormone-inducible transmembrane protein
gene1897	F>L	L>F	L>F	chromodomain-helicase-DNA-binding protein 1 isoform X1
gene1914	F>L	L>F	L>F	high affinity cAMP-specific and IBMX-insensitive 3',5'-cyclic
gene2098	F>L	L>F	L>F	glutamate receptor 1
gene2186	F>L	L>F	L>F	putative oxidoreductase GLYR1 homolog
gene2693	F>L	L>F	L>F	monocarboxylate transporter 12
gene2928	F>L	L>F	L>F	splicing factor 3B subunit 5
gene3442	F>L	L>F	L>F	programmed cell death protein 2-like
gene3459	F>L	L>F	L>F	coronin-2B
gene3460	F>L	L>F	L>F	NEDD8-activating enzyme E1 catalytic subunit
gene3463	F>L	L>F	L>F	exocyst complex component 6B
gene38	F>L	L>F	L>F	extra macrochaetae
gene4105	F>L	L>F	L>F	hemicentin-2
gene4239	F>L	L>F	L>F	growth hormone-regulated TBC protein 1-A
gene4490	F>L	L>F	L>F	uncharacterized LOC113218777
gene463	F>L	L>F	L>F	synaptotagmin-4
gene5300	F>L	L>F	L>F	syntaxin-12
gene5302	F>L	L>F	L>F	vacuolar H+ ATP synthase 16 kDa proteolipid subunit
gene5404	F>L	L>F	L>F	uncharacterized LOC724221
gene5450	F>L	L>F	L>F	E3 ubiquitin-protein ligase RNF126
gene559	F>L	L>F	L>F	U3 small nucleolar RNA-associated protein 6 homolog
gene5763	F>L	L>F	L>F	coiled-coil domain-containing protein 112
gene5793	F>L	L>F	L>F	14-3-3 protein epsilon
gene6010	F>L	L>F	L>F	methylosome protein 50
gene6122	F>L	L>F	L>F	uncharacterized LOC726661
gene6155	F>L	L>F	L>F	histone H2B
gene6218	F>L	L>F	L>F	uncharacterized LOC408909
gene6281	F>L	L>F	L>F	uncharacterized LOC413002

(Table B.5 continued)

gene6407	F>L	L>F	L>F	ATP synthase lipid-binding protein, mitochondrial
gene6447	F>L	L>F	L>F	NADH dehydrogenase [ubiquinone] 1 alpha subcomplex subunit 10, mitochondrial
gene7932	F>L	L>F	L>F	histone demethylase UTY
gene8288	F>L	L>F	L>F	mRNA turnover protein 4 homolog
gene9945	F>L	L>F	L>F	tubulin alpha-1 chain
gene10888	L>F	F>L	F>L	E3 ubiquitin-protein ligase RING1
gene11389	L>F	F>L	F>L	protein phosphatase 1 regulatory subunit 3C-B
gene1373	L>F	F>L	F>L	mucin-12 isoform X1
gene3849	L>F	F>L	F>L	signal-induced proliferation-associated 1-like protein 2
gene3949	L>F	F>L	F>L	protein croquemort
gene4523	L>F	F>L	F>L	uncharacterized protein LOC725128
gene7042	L>F	F>L	F>L	muscarinic acetylcholine receptor DM1
gene790	L>F	F>L	F>L	uncharacterized LOC100578657
gene8541	L>F	F>L	F>L	F-actin-monooxygenase Mical
gene9457	L>F	F>L	F>L	prohormone-1
gene9936	L>F	F>L	F>L	CDK5RAP1-like protein

References

Alaux, C., Le Conte, Y., Adams, H. A., Rodriguez-Zas, S., Grozinger, C. M., Sinha, S. and Robinson, G. E. (2009). Regulation of brain gene expression in honey bees by brood pheromone. *Genes, Brain Behav.* 8, 309–319.

Grozinger, C. M., Fan, Y., Hoover, S. E. R. and Winston, M. L. (2007). Genome-wide analysis reveals differences in brain gene expression patterns associated with caste and reproductive status in honey bees (*Apis mellifera*). *Mol. Ecol.* 16, 4837–4848.

Harpur, B. A., Kent, C. F., Molodtsova, D., Lebon, J. M. D., Alqarni, A. S., Owayss, A. A., and Zayed, A. (2014). Population genomics of the honey bee reveals strong signatures of positive selection on worker traits. *Proc. Natl. Acad. Sci.* 111, 2614–2619.

Jones, B. M., Kingwell, C. J., Wcislo, W. T. and Robinson, G. E. (2017). Caste-biased gene expression in a facultatively eusocial bee suggests a role for genetic accommodation in the evolution of eusociality. *Proc. R. Soc. B Biol. Sci.* 284, 20162228.

Kapheim, K. M., Pan, H., Li, C., Salzberg, S. L., Puiu, D., Magoc, T., Robertson, H.M., Hudson, M.E., Venkat, A., Fischman, B.J. et al. (2015). Genomic signatures of evolutionary transitions from solitary to group living. *Science* 348, 1139–1143.

Marshall, H., Lonsdale, Z. N. and Mallon, E. B. (2019). Methylation and gene expression differences between reproductive castes of bumblebee workers. *bioRxiv*, 517698.

Toth, A. L., Várala, K., Henshaw, M. T., Rodriguez-Zas, S. L., Hudson, M. E. and Robinson, G. E. (2010). Brain transcriptomic analysis in paper wasps identifies genes associated with behaviour across social insect lineages. *Proc. R. Soc. B Biol. Sci.* 277, 2139–2148.

Wheeler, M. M., Ament, S. A., Rodriguez-Zas, S. L. and Robinson, G. E. (2013). Brain gene expression changes elicited by peripheral vitellogenin knockdown in the honey bee. *Insect Mol. Biol.* 22, 562–573.

Whitfield, C. W., Cziko, A. M. and Robinson, G. E. (2003). Gene expression profiles in the brain predict behavior in individual honey bees. *Science* 302, 296–299.

Woodard, S. H., Fischman, B. J., Venkat, A., Hudson, M. E., Varala, K., Cameron, S. A., Clark, A. G. and Robinson, G. E. (2011). Genes involved in convergent evolution of eusociality in bees. *Proc. Natl. Acad. Sci.* 108, 7472–7477.

APPENDIX C: SUPPLEMENTARY DATASETS

Dataset C.1. Lists of differentially expressed genes (DEGs) for each pairwise comparison of egg-layers (L), foragers (F), and generalists (G).

Dataset C.2. Lists of differentially accessible chromatin region (DAR) and peak (DAP) results for each pairwise comparison of egg-layers (L), foragers (F), and generalists (G).

Dataset C.3. Gene Ontology (GO) enrichment results for all DEG lists for each pairwise comparison of egg-layers (L), foragers (F), and generalists (G), as well as for extreme loading genes of principal components (PCs) correlated with behavioral variation.

Dataset C.4. Detailed information about studies used for comparison with F vs. L DEGs. Gene lists from each study, as well as overlapping gene lists for each comparison and results of representation factor analysis, are provided. Conversion lists between different *Apis mellifera* annotation versions and reciprocal best BLASTP results between each pairs of species used for comparison are also reported.

Dataset C.5. Gene Ontology (GO) enrichment results for all DAR and DAP lists, as well as for extreme loadings genes of principal components (PCs) correlated with behavioral variation.